

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search using frame_plus_p2n model

Run on: May 9, 2006, 13:15:42 ; Search time 458.2 Seconds

(without alignments)
3163.483 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17

Perfect score: 81

Sequence: 1 MALEVLMLLAVALIWTGA 17

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 588141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q=/abs/ABSWEB.spool/US10664025/runat_08052006_173641_29351/app.query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss03p
-USER=US10664025.@CGN_1_17724_@runat_08052006_173641_29351 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sta.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_hlg.*

15: gb_pl.*

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	469	6	BD109279
2	81	100.0	469	6	AR413726
3	81	100.0	469	6	AX970560

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

4	81	100.0	470	6	CS072276
5	81	100.0	477	6	CQ737424
6	81	100.0	457	6	AX588689
7	81	100.0	500	6	AX588688
8	81	100.0	512	6	BD107959
9	81	100.0	512	6	AR412406
10	81	100.0	512	6	AX969240
11	81	100.0	560	6	BD110503
12	81	100.0	560	6	AR414950
13	81	100.0	560	6	AX971784
14	81	100.0	642	6	BD109278
15	81	100.0	642	6	AR413725
16	81	100.0	642	6	AX970559
17	81	100.0	691	6	CS072271
18	81	100.0	1337	6	AX574405
19	81	100.0	1480	6	AX600204
20	81	100.0	1492	6	AX748179
21	81	100.0	1492	8	AK093517
22	81	100.0	1502	8	BC036256
23	81	100.0	1646	8	BC048121
24	81	100.0	1898	6	BD275948
25	81	100.0	10115	6	AX574404
26	81	100.0	100317	14	AP000409
27	81	100.0	101031	14	AP000622
28	81	100.0	176944	8	AP000790
29	81	100.0	188367	8	AC013807
30	62	76.5	655	9	BC099498
31	62	76.5	132494	9	AC126036
32	62	76.5	203047	9	AC127289
33	54	66.7	372	6	BD119756
34	54	66.7	372	6	AR424203
35	54	66.7	372	6	AX984897
36	53	65.4	745	2	TSA430566
37	52	64.2	105951	15	AC141106
38	52	64.2	183851	9	AC139357
39	52	64.2	185548	9	AC005743
40	52	64.2	193347	9	AC157365
41	51	63.0	110000	1	BA000037
42	51	63.0	178865	14	AC133135
43	51	63.0	288704	14	AC163368
44	51	63.0	300169	1	AE016801
45	50	61.7	11443	1	AE000823

ALIGNMENTS

RESULT 1

BD109279

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD109279 469 bp DNA linear PAT 18-SEP-2002
EST and encoded human protein.

BD109279.1 GI:23204097

JP 2002010789-A/1356

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 469)

Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.

EST and encoded human protein

Patent: JP 2002010789-A 1356 15-JAN-2002;

GENSET CORP

OS Homo sapiens (human)

PN JP 2002010789-A/1356

PD 15-JAN-2002

PF 07-AUG-2000 JP 2002080989

PR 05-AUG-1999 US 60/147499

PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE

GIORDANO

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

C12N1/21

PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC

ORIGIN		GENSET (FR)		Location/Qualifiers	
Alignment Scores:		source		1..497	
Pred. No.:	6.65e-06	Length:		470	
Score:	81.00	Matches:		17	
Percent Similarity:	100.0%	Conservative:		0	
Best Local Similarity:	100.0%	Mismatches:		0	
Query Match:	100.0%	Indels:		0	
DB:	6	Gaps:		0	
US-10-664-025-3903_COPY_1_17 (1-17) x CS072276 (1-470)					
Qy		1		MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17	
Db		137		ATGGCGTTAGAGTCCTTGATGCTCCCTCGCTGCTTGATTGGACCGGTCT 187	
RESULT 5					
CQ737424		477 bp		DNA linear	
LOCUS		Sequence 23358 from Patent WO02068579.		PAT 03-FEB-2004	
DEFINITION		CQ737424			
ACCESSION		CQ737424.1		GI:42335713	
VERSION					
KEYWORDS		Homo sapiens (human)			
SOURCE		Homo sapiens			
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE		1		Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	
AUTHORS				Kits, such as nucleic acid arrays, comprising a majority of	
TITLE				humanexons or transcripts, for detecting expression and other uses	
JOURNAL				thereof	
PATENT				Patent: WO 02068579-A 23358 06-SEP-2002;	
PE Corporation (NY) (US)				PE Corporation (NY) (US)	
FEATURES		Location/Qualifiers			
source		1..477			
/organism="Homo sapiens"					
/mol_type="unassigned DNA"					
/db_xref="taxon:9606"					
ORIGIN					
Alignment Scores:		6.76e-06		Length: 477	
Pred. No.:	81.00	Matches:		17	
Score:	100.0%	Conservative:		0	
Percent Similarity:	100.0%	Mismatches:		0	
Best Local Similarity:	100.0%	Indels:		0	
Query Match:	100.0%	Gaps:		0	
US-10-664-025-3903_COPY_1_17 (1-17) x CQ737424 (1-477)					
Qy		1		MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17	
Db		1		ATGGCGTTAGAGTCCTTGATGCTCCCTCGCTGCTTGATTGGACCGGTCT 51	
RESULT 6					
AX588689		497 bp		DNA linear	
LOCUS		Sequence 564 from Patent WO02083898.		PAT 24-JAN-2003	
DEFINITION		AX588689			
ACCESSION		AX588689.1		GI:27900346	
VERSION					
KEYWORDS		Homo sapiens (human)			
SOURCE		Homo sapiens			
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE		1		Bejanin, S., Tanaka, H., Dumas Milne Edwards, J.B., Jobert, S. and	
AUTHORS				Giordano, J.Y.	
TITLE				Full-length human cdnas encoding potentially secreted proteins	
JOURNAL				Patent: WO 02083898-A 564 24-OCT-2002;	
FEATURES					
source		1..500			
/organism="Homo sapiens"					
/mol_type="unassigned DNA"					
/db_xref="taxon:9606"					
/note="unassigned protein product"					
/codon_start=1					
/protein_id="CAD61627.1"					
/db_xref="GI:27900345"					
/translation="MALEVLMLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNL					
YIPADELHLGMCPCANRIHTYVYEFILVVDCCGIRTVVSEETLLFQTELYFTPRNID					
HDPQEIHLCESTSRKSVWLTPVSTENEIKLDPSFPFIADPQT"					
72..122					
sig_peptide					
/note="Von Heijne matrix score 10.69999998092651 seq					
VLMMLAVLIWTGA/EN"					
ORIGIN					
Alignment Scores:		7.11e-06		Length: 500	
Pred. No.:	81.00	Matches:		17	
Score:	100.0%	Conservative:		0	
Percent Similarity:	100.0%				

```

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x AX588688 (1-500)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuIleTrpThrGlyAla 17
|||||
Db 72 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTGACCGGTGCT 122

RESULT 8
BD107959 512 bp DNA linear PAT 18-SEP-2002
LOCUS EST and encoded human protein.
DEFINITION BD107959
ACCESSION BD107959
VERSION BD107959.1 GI:23202777
KEYWORDS JP 2002010789-A/36.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 512)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 36 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/36
PD 15-JAN-2002
PF 07-AUG-2000 JP 2002080989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 10.6999998092651
CC seq VMLMLAVLIWTGA/EN
FH Key Location/Qualifiers
FT CDS 53..511
FT sig peptide 53..103.
FEATURES
source 1..512
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 7.29e-06 Length: 512
Pred. No.: 81.00 Matches: 17
Score: 81.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-664-025-3903_COPY_1_17 (1-17) x BD107959 (1-512)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTGACCGGTGCT 103

RESULT 9
AR412406 512 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 43 from patent US 6639063.
DEFINITION AR412406
ACCESSION AR412406
VERSION AR412406.1 GI:40167516
KEYWORDS
SOURCE Unknown.

```

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 512)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 43 28-OCT-2003;
GENSET S.A.;;
WOX;
FEATURES
source 1..512
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 7.29e-06 Length: 512
Pred. No.: 81.00 Matches: 17
Score: 81.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-664-025-3903_COPY_1_17 (1-17) x AR412406 (1-512)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTGACCGGTGCT 103

RESULT 10
AX969240 512 bp DNA linear PAT 15-JAN-2004
LOCUS Sequence 43 from Patent EP1104808.
DEFINITION AX969240
ACCESSION AX969240
VERSION AX969240.1 GI:40975399
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE ESTs and encoded human proteins
JOURNAL Patent: EP 1104808-A 43 06-JUN-2001;
GENSET (FR)
FEATURES
source 1..512
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
53..>511
/note="unnamed protein product"

CDS
/codon_start=1
/protein_id="CAF14348.1"
/db_xref="GI:40975400"
/translation="MALEVLMLLAVLIWTGAENLHVKISCSLDMLMSVIPVAESRNL
YIFDELHLGMGCPANRTHTVYVFILVDRCGIRTVVSEETLLFOTELYFTPRNID
HDPQEIHLCESTSRKSVWLTPVSTENEIKLDPESFIADFTTAELGLLIF"
53..103
sig_peptide
/note="Von Heijne matrix score 10.6999998092651 seq
VLMMLAVLIWTGA/EN"

ORIGIN
Alignment Scores: 7.29e-06 Length: 512
Pred. No.: 81.00 Matches: 17
Score: 81.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-664-025-3903_COPY_1_17 (1-17) x AX969240 (1-512)

```


Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCT 103

RESULT 11
LOCUS BD110503 560 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD110503
VERSION BD110503.1 GI:23205321
KEYWORDS JP 2002010789-A/2580.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 560)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 2580 15-JAN-2002;
GENSET CORP

COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/2580
PD 15-JAN-2002
PF 07-AUG-2000 JP 20020280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI
GIORDANO
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS 372..539.

FEATURES
source 1..560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 8.03e-06 Length: 560
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BD110503 (1-560)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCT 103

RESULT 12
LOCUS AR414950 560 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2587 from patent US 6639063.
ACCESSION AR414950
VERSION AR414950.1 GI:40170060
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 560)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 2587 28-OCT-2003;
Genset S.A.;;
WOX; Location/Qualifiers

FEATURES
source 1..560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 8.03e-06 Length: 560
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BD110503 (1-560)

source 1..560
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 8.03e-06 Length: 560
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x AR414950 (1-560)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCT 103

RESULT 13
LOCUS AX971784 560 bp DNA linear PAT 15-JAN-2004
DEFINITION Sequence 2587 from Patent EP1104808.
ACCESSION AX971784
VERSION AX971784.1 GI:40980487
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE ESTs and encoded human proteins
JOURNAL Patent: EP 1104808-A 2587 06-JUN-2001;
Genset (FR)

FEATURES
source 1..560
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
372..>539
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF16892.1"
/db_xref="GI:40980488"
/translation="MTLRKSIWECSTSRKSVMLTPVSTENEIKLDPSPTADPQTAE
ELGLLSSEPNLL"

ORIGIN
Alignment Scores:
Pred. No.: 8.03e-06 Length: 560
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x AX971784 (1-560)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCT 103

RESULT 14
LOCUS BD109278 642 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD109278
VERSION BD109278.1 GI:23204096
KEYWORDS JP 2002010789-A/1355.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 642)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein

JOURNAL Patent: JP 2002010789-A 1355 15-JAN-2002;
GENSET CORP

COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/1355
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 10.699998092651
CC seq VMLLLAVLIWTGA/EN
FH key Location/Qualifiers
FT CDS 53..526
FT sig peptide 53..103.
FT Location/Qualifiers

FEATURES
source 1..642
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9.3e-06 Length: 642
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BD109278 (1-642)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
DB 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTGGACCGGTGCT 103

RESULT 15
AR413725
LOCUS AR413725 642 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1362 from patent US 6639063.
ACCESSION AR413725
VERSION AR413725.1 GI:40168835
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 642)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 1362 28-OCT-2003;
Genset S.A.;;
WOX;

FEATURES
source Location/Qualifiers
1..642
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 9.3e-06 Length: 642
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-664-025-3903_COPY_1_17 (1-17) x AR413725 (1-642)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
DB 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTGGACCGGTGCT 103

Search completed: May 9, 2006, 14:49:46
Job time : 460.2 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 13:15:10 ; Search time 88.4 Seconds
(without alignments)
1922.507 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17
Perfect score: 81
Sequence: 1 MALEVLMLLAVLITGCA 17

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10664025/runat_08052006_173639_29334/app_query.fasta 1
-DB=N Geneseq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss07
-USER=US10664025 @CGN_1_1_900 @runat_08052006_173639_29334 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	470	5 AAh64748	Aah64748 Human sec
2	81	100.0	497	8 AB236691	Abz36691 Human GEN
3	81	100.0	500	8 AB236690	Abz36690 Human GEN
4	81	100.0	691	5 AAh64743	Aah64743 Human sec

5	81	100.0	1337	6 ABK90053	Abk90053 DNA encod
6	81	100.0	1480	6 ABS78646	Abs78646 Human CDN
7	81	100.0	1492	10 ADB63550	Adb63550 Human CDN
8	81	100.0	1898	3 AAC69528	Aac69528 Human sec
9	81	100.0	1898	8 ACC50642	Acc50642 Human sec
10	81	100.0	1898	8 AB271354	Abz71354 Secreted
11	81	100.0	1898	9 ADB91291	Adb91291 Human sec
12	81	100.0	1898	10 ADC73723	Adc73723 Human sec
13	81	100.0	10115	6 ABK90052	Abk90052 Human GSS
c 14	52	64.2	185548	13 ADV34986	Adv34986 Murine cD
c 15	49	60.5	34886	12 ADO34435_6	Continuation (7 of
c 16	47	58.0	65	3 AAK99161	Aak99161 65-mer he
c 17	47	58.0	696	13 ADR62236	Adr62236 Cotton cD
c 18	47	58.0	3407	4 ABL18088	Ab18088 Drosophil
c 19	46	56.8	737	6 ABQ19936	Abq19936 Oligonuc
c 20	46	56.8	737	6 ABQ19937	Abq19937 Oligonuc
c 21	46	56.8	2031	12 ADN11104	Adn11104 Vaccinia
c 22	46	56.8	2459	4 ABL17617	Ab17617 Drosophil
c 23	46	56.8	3312	10 ADE58050	Ades8050 Rat Gene
c 24	46	56.8	20460	4 ABL17616	Ab17616 Drosophil
c 25	46	56.8	20460	4 ABL20770	Ab120770 Drosophil
c 26	46	56.8	349980	6 ABQ81845	Abq81845 Bifidobac
c 27	46	56.8	349980	6 ABQ81844	Abq81844 Bifidobac
c 28	45	55.6	243	6 ABV86822	Abv86822 Human col
c 29	45	55.6	369	5 AAS33650	Aas33650 Human CDN
c 30	45	55.6	492	11 ABD05447	Abd05447 Pseudomon
c 31	45	55.6	500	12 ACH75891	Ach75891 Human gen
c 32	45	55.6	507	13 ACN58117	Acn58117 Cotton gy
c 33	45	55.6	590	12 ACH89603	Ach89603 Human gen
c 34	45	55.6	596	6 ABL92377	Ab192377 Human sec
c 35	45	55.6	597	13 ADQ53491	Adq53491 Novel can
c 36	45	55.6	682	3 AAF14449	Aaf14449 Aspergill
c 37	45	55.6	682	13 ADU58490	Adu58490 Aspergill
c 38	45	55.6	682	14 ADZ96493	Adz96493 Aspergill
c 39	45	55.6	965	5 AAS34723	Aas34723 Human DNA
c 40	45	55.6	965	5 AAS34722	Aas34722 Human DNA
c 41	45	55.6	973	6 AB192353	Ab192353 Human sec
c 42	45	55.6	1003	4 AAH32256	Aah32256 Human oif
c 43	45	55.6	1077	11 ABD05379	Abd05379 Pseudomon
c 44	45	55.6	1278	11 ABD05292	Abd05292 Pseudomon
c 45	45	55.6	1713	12 ADQ03085	Adq03085 P. aerugi

ALIGNMENTS

RESULT 1
AAH64748
ID AAh64748 standard; cDNA; 470 BP.
XX
AC AAh64748;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 24.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
XX Homo sapiens.
XX
XX WO200142451-A2.
XX
PD 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-IB001938.
XX
XX 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX
(GEST) GENSET.
XX
Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX

DR WPI; 2001-367870/38.
XX P-PSDB; AAG89145.
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
PS Claim 7; Page 583; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX
SQ Sequence 470 BP; 137 A; 87 C; 107 G; 139 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.00015 Length: 470
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x AAH64748 (1-470)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 137 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTGGACCGTGCT 187

RESULT 2
ABZ36691
ID ABZ36691 standard; cDNA; 497 BP.
XX
AC ABZ36691;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET coding sequence SEQ ID 564.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2001; 2001WO-IB000914.
XX
PR 18-APR-2001; 2001WO-IB000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX

DR WPI; 2003-075548/07.
XX
PT New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
PS Claim 12; Page 579; 735pp; English.
XX
CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps
XX
SQ Sequence 497 BP; 127 A; 99 C; 117 G; 154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.00016 Length: 497
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ABZ36691 (1-497)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 52 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTGGACCGTGCT 102

RESULT 3
ABZ36690
ID ABZ36690 standard; cDNA; 500 BP.
XX
AC ABZ36690;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET coding sequence SEQ ID 563.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2001; 2001WO-IB000914.
XX
PR 18-APR-2001; 2001WO-IB000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX
DR WPI; 2003-075548/07.
XX
PT New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
PS Claim 12; Page 578; 735pp; English.
XX

CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps
XX
SQ Sequence 500 BP; 129 A; 100 C; 113 G; 155 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 0.000161 Length: 500
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ABZ36690 (1-500)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 72 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTGGACCGGTGCT 122

RESULT 4

AAH64743

ID AAH64743 standard; cDNA; 691 BP.

XX AC AAH64743;

XX DT 11-SEP-2001 (first entry)

XX DE Human secreted protein cDNA, SEQ ID NO: 19.

XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX KW GENSET; ss.

XX OS Homo sapiens.

XX PN WO200142451-A2.

XX PD 14-JUN-2001.

XX PF 07-DEC-2000; 2000WO-IB001938.

XX PR 08-DEC-1999; 99US-0169629P.

XX PR 06-MAR-2000; 2000US-0187470P.

XX XX (GEST) GENSET.

XX PA Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX PI WPI: 2001-367870/38.

XX DR P-PSDB; AAG89140.

XX PT Full length GENSET human nucleic acids encoding potentially secreted
XX PT proteins, useful in gene therapy and vaccination against a variety of
XX PT diseases, and for diagnosis of those diseases.

XX PS Claim 7; Page 579; 921pp; English.

XX CC The invention relates to full length GENSET human nucleic acids encoding
XX CC potentially secreted proteins. The nucleic acids and the polypeptides
XX CC they encode may be used in the prevention, treatment and diagnosis of
XX CC diseases associated with inappropriate GENSET gene expression. For
XX CC example, they be used to treat disorders associated with decreased GENSET
XX CC gene expression by rectifying mutations or deletions in a patient's
XX CC genome that affect the activity of GENSET or by supplementing the
XX CC patients own production of GENSET polypeptides. Conversely, antisense
XX CC nucleic acid molecules may be administered to down regulate GENSET
XX CC expression by binding with the cells' own genes and preventing their

CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX
SQ Sequence 691 BP; 198 A; 126 C; 151 G; 216 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.000238 Length: 691
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x AAH64743 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTGGACCGGTGCT 92

RESULT 5

ABK90053

ID ABK90053 standard; DNA; 1337 BP.

XX AC ABK90053;

XX DT 05-NOV-2002 (first entry)

XX DE DNA encoding human GSSP3 polypeptide.

XX KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
XX KW body mass; serum glucose regulation; body weight loss; obesity;
XX KW metabolic-related disorder; impaired glucose tolerance; stroke;
XX KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
XX KW hypertension; syndrome C; type I diabetes; type II diabetes;
XX KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
XX KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
XX KW lipatrophy; physical performance; exercise; dyslexia; schizophrenia;
XX KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
XX KW gene; ds.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX CDS 53..529

XX FT /*tag= a

XX FT /product= "GSSP3 polypeptide"

XX FT 53..103

XX FT /*tag= b

XX FT 104..526

XX FT /*tag= c

XX XX WO200260466-A2.

XX PN

XX PD 08-AUG-2002.

XX XX 01-FEB-2002; 2002WO-IB001333.

XX XX 02-FEB-2001; 2001US-0266156P.

XX PA (GEST) GENSET.

XX PI Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;

XX PI Bihain B;

XX DR WPI: 2002-608487/65.

XX DR P-PSDB; ABG31324.

XX XX

PT Reducing circulating glucose levels or increasing insulin sensitivity,
PT useful for reducing body mass or preventing body weight gain, comprises
PT administering composition comprising GSSP3 polypeptide.
XX
PS Disclosure; Page 95-96; 97pp; English.
XX
CC The present invention relates to the isolation of human GSSP3
CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
CC polypeptide reduces circulating blood glucose levels, increases insulin
CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
CC polynucleotide sequences are useful in serum glucose regulation, fatty
CC acid metabolism, body weight loss, and prevention of body weight gain.
CC Compositions comprising GSSP3 polypeptides are useful for controlling
CC blood glucose levels, for treating metabolic-related diseases or
CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
CC syndrome C, type I or II diabetes, diabetes related complications,
CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
CC acanthosis nigricans, leprechaunism and lipotrophy). The polypeptides
CC are also useful to improve physical performance during work or exercise,
CC and to treat dyslexia, attention-deficit disorder, attention-
CC deficit/hyperactivity disorder, and psychiatric disorders such as
CC schizophrenia. The present sequence encodes human GSSP3 polypeptide
XX
SQ Sequence 1337 BP; 457 A; 214 C; 240 G; 426 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.000524 Length: 1337
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ABK90053 (1-1337)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 53 ATGGCGTTAGAGCTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCT 103
|||||

RESULT 6
ABS78646
ID ABS78646 standard; cDNA; 1480 BP.
XX
AC ABS78646;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding, CGDD4, INCYTE 5284076CB1.
XX
KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200272830-A2.
XX
PD 19-SEP-2002.
XX
PP 08-FEB-2002; 2002WO-US003715.
XX
PR 09-FEB-2001; 2001US-0268111P.

PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
DR WPI; 2002-723356/78.
XX
DR P-PSDB; ABG97353.
XX
PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
XX Claim 5; Page 169; 181pp; English.
XX
PS The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotides and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 1480 BP; 463 A; 244 C; 265 G; 508 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.000591 Length: 1480
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ABS78646 (1-1480)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
|||||
Db 29 ATGGCGTTAGAGCTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCT 79
|||||

RESULT 7
ADB63550
ID ADB63550 standard; cDNA; 1492 BP.
XX
AC ADB63550;
XX
DT 04-DEC-2003 (first entry)

XX DE Human cDNA encoding clone TESTI20282420.
XX KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW tissue regeneration; cell regeneration; membrane protein;
KW signal transduction-related protein; transcription-related protein;
KW osteoporosis; neurological disease; cancer; tumour.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 52..528
FT /*tag= a
FT /product= "Clone TESTI20282420 protein"
XX PN EP1308459-A2.
XX PD 07-MAY-2003.
XX PF 28-MAR-2002; 2002EP-00007401.
XX PR 05-NOV-2001; 2001JP-00379298.
XX PR 25-JAN-2002; 2002US-00350978.
XX (HELI-) HELIX RES INST.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI: 2003-450961/43.
XX P-PSDB; ADB65520.
XX PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX Claim 1; Page: 222pp; English.
XX The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesising the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX SQ Sequence 1492 BP; 452 A; 250 C; 272 G; 518 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.000597	Length:	1492
Score:	81.00	Matches:	17
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-664-025-3903_COPY_1_17 (1-17) x ADB63550 (1-1492)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

Db 52 ATGGCGTTAGAGCTCTTGATGCTCTCGCTGCTCTGATTGGACCGTGCT 102

RESULT 8

AAC69528

ID AAC69528 standard; DNA; 1898 BP.

XX AC AAC69528;

XX DT 31-JAN-2001 (first entry)

XX XX Human secreted protein gene 17 clone HTELS08.

XX XX Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;

KW cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;

KW nontropic; antibacterial; virucide; neurotoxic; ophthalmological;

KW vulnery; gene therapy; infection; human; secreted protein; ss.

XX OS Homo sapiens.

XX PN WO200061623-A1.

XX PD 19-OCT-2000.

XX PF 06-APR-2000; 2000WO-US008979.

XX PR 09-APR-1999; 99US-0128693P.

XX PR 26-APR-1999; 99US-0130991P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Ruben SM, Ni J, Konatsoulis GA, Rosen CA, Soppet DR, Shi Y;

PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;

PI Young PE;

XX WPI: 2000-647418/62.

XX New nucleic acid molecules encoding 62 human secreted proteins for
CC diagnosing, preventing, treating or ameliorating medical conditions and
CC used as food additives or preservatives.
XX Claim 1; Page 543-544; 716pp; English.
XX The invention relates to the isolation of genes (AAC69512-C69587)
CC encoding 62 human secreted proteins (AAB38321-B38396). The genes can be
CC used to generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (AAC69503) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of
CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;
CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angiogenesis;
CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections
CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.
CC corneal infection. The polypeptides can also be used to aid wound healing
CC and epithelial cell proliferation, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis

SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	0.000796	Length:	1898
Score:	81.00	Matches:	17
Percent Similarity:	100.0%	Conservative:	0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x AAC69528 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla 17
DB 15 ATGGCGTTAGAACTCTTGATGCTCCTCGTGTCTTGATTGACCGGTGCT 65

RESULT 9
ACCS0642
ID ACCS0642 standard; cDNA; 1898 BP.
XX
AC ACCS0642;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein coding sequence, SEQ ID 309.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
KW vulnery; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002WO-US009785.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-129429/12.
XX
DR Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
PT disorders such as arrhythmia.
XX
PS Claim 21; SEQ ID NO 309; 1881pp; English.
XX
CC The present invention relates to novel human secreted proteins (ABR47633-
CC ABR48145) and their coding sequences (ACCS0344-ACCS0856). The proteins
CC and their coding sequences are useful for the preparation of a diagnostic
CC or pharmaceutical composition for diagnosing or treating a cardiovascular
CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
CC arteriosclerosis and myocardial ischaemia), neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in
CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 0.000796 Length: 1898
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ACCS0642 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla 17
DB 15 ATGGCGTTAGAACTCTTGATGCTCCTCGTGTCTTGATTGACCGGTGCT 65

RESULT 10
ABZ71354
ID ABZ71354 standard; cDNA; 1898 BP.
XX
AC ABZ71354;
XX
DT 04-APR-2003 (first entry)
XX
DE Secreted protein-encoding gene 165 cDNA clone HTELS08, SEQ ID NO:175.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulnery; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200276488-A1.
XX
PD 03-OCT-2002.
XX
PF 19-MAR-2002; 2002WO-US008276.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
XX P-PSDB; ABR00175.
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers.
XX
PS Claim 21; Page 853-854; 1216pp; English.
XX
CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome

CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein-
CC encoding cDNA clone of the invention

XX SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 0.000796 Length: 1898
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ABZ71354 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
Db 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTGCTGCTGCTGCTGCTGCT 65

RESULT 11

ADB91291
ID ADB91291 standard; cDNA; 1898 BP.

XX ADB91291;

XX 04-DEC-2003 (first entry)

XX Human secreted protein cDNA #SEQ ID 237.

XX Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.

XX Homo sapiens.

XX WO2003004622-A2.

XX 16-JAN-2003.

XX 19-MAR-2002; 2002WO-US008124.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-229407/22.

XX Nucleic acid encoding a human secreted protein is useful in diagnosing or

XX treating diabetes or conditions related to diabetes.

XX Claim 9; SEQ ID NO 237; 1537pp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-
CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
CC ADB91834. Also disclosed is a recombinant vector comprising a
CC polynucleotide of the invention, and a recombinant host cell comprising
CC the recombinant vector. The polypeptide of the invention is useful in
CC identifying a binding partner by contacting the polypeptide with a
CC binding partner, and determining whether the binding partner increases or
CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
CC antibody or its fragment, agonist or antagonist are useful for preparing
CC a pharmaceutical composition for diagnosing or treating diabetes or
CC conditions related to diabetes. The present sequence is that of the human
CC immunoglobulin Fc portion used to generate fusion proteins, increasing
CC the stability of the fused protein as compared to the secreted protein
CC only. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pet_sequences.

XX

SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 0.000796 Length: 1898
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ADB91291 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
Db 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTGCTGCTGCTGCTGCTGCT 65

RESULT 12

ADC73723

ID ADC73723 standard; DNA; 1898 BP.

XX ADC73723;

XX 01-JAN-2004 (first entry)

XX Human secreted protein-related DNA - SEQ ID 356.

XX antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
KW antidiabetic; immunosuppressive; dermatological; nephrotropic;
KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;
KW haemopoietic; haematologic; anaemia; autoimmune disorder;
KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
KW human; gene; ds.

XX Homo sapiens.

XX WO2003038063-A2.

XX 08-MAY-2003.

XX 19-MAR-2002; 2002WO-US008277.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-430516/40.

XX P-PSDB; ADC74338.

XX New human secreted polypeptide for diagnosing, preventing or treating

XX hemopoietic or hematologic disorders (e.g. anemia), autoimmune

XX disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or

XX atherosclerosis).

XX Claim 27; SEQ ID NO 356; 2272pp; English.

XX The invention relates to a novel human secreted polypeptide comprising a
CC defined sequence given in the specification. The polypeptide, nucleic
CC acid molecule, antibody, agonist or antagonist of the invention may be
CC useful for preparing a composition for diagnosing or treating a
CC haemopoietic or haematologic disorder such as anaemia, autoimmune
CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
CC diabetes, systemic lupus erythematosus or glomerulonephritis,
CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
CC disease, wounds and hyperproliferative disorders including
CC atherosclerosis or cancer, as well as bacterial, viral, fungal or

CC parasitic infections. The polypeptide may also be used during gene
 CC therapy procedures and for identifying a binding partner by contacting
 CC the polypeptide with a binding partner and determining whether the
 CC binding partner increases or decreases the activity of the polypeptide.
 CC The current sequence is that of the human secreted protein-related DNA of
 CC the invention.

XX Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:
 Pred. No.: 0.000796 Length: 1898
 Score: 81.00 Matches: 17
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ADC73723 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleThrThrGlyAla 17
 Db 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGTGCT 65

RESULT 13

ABK90052

ID ABK90052 standard; DNA; 10115 BP.

XX AC ABK90052;

XX DT 05-NOV-2002 (first entry)

XX DE Human GSSP3 genomic sequence.

XX Human; GSSP3; circulating blood glucose level; insulin sensitivity;
 KW body mass; serum glucose regulation; body weight loss; obesity;
 KW metabolic-related disorder; impaired glucose tolerance; stroke;
 KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
 KW hypertension; syndrome C; type I diabetes; type II diabetes;
 KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
 KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
 KW lipotrophy; physical performance; exercise; dyslexia; schizophrenia;
 KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
 KW gene; db.

XX OS Homo sapiens.

XX PN WO200260466-A2.

XX PD 08-AUG-2002.

XX PF 01-FEB-2002; 2002WO-IB001333.

XX PR 02-FEB-2001; 2001US-0266156P.

XX PA (GEST) GENSET.

XX PI Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
 PI Bihain B;

XX WIPI; 2002-608487/65.

XX Reducing circulating glucose levels or increasing insulin sensitivity,
 PT useful for reducing body mass or preventing body weight gain, comprises
 PT administering composition comprising GSSP3 polypeptide.

PS Disclosure; Page 90-95; 97pp; English.

XX The present invention relates to the isolation of human GSSP3
 CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
 CC polypeptide reduces circulating blood glucose levels, increases insulin
 CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
 CC polynucleotide sequences are useful in serum glucose regulation, fatty
 CC acid metabolism, body weight loss, and prevention of body weight gain.

CC Compositions comprising GSSP3 polypeptides are useful for controlling
 CC blood glucose levels, for treating metabolic-related diseases or
 CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
 CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
 CC syndrome C, type I or II diabetes, diabetes related complications,
 CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
 CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
 CC acanthosis nigricans, leprechaunism and lipotrophy). The polypeptides
 CC are also useful to improve physical performance during work or exercise,
 CC and to treat dyslexia, attention-deficit disorder, attention-
 CC deficit/hyperactivity disorder, and psychiatric disorders such as
 CC schizophrenia. The present sequence encodes human GSSP3 polypeptide
 XX Sequence 10115 BP; 2860 A; 1941 C; 1815 G; 3475 T; 0 U; 24 Other;

Alignment Scores:
 Pred. No.: 0.00589 Length: 10115
 Score: 81.00 Matches: 17
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x ABK90052 (1-10115)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleThrThrGlyAla 17
 Db 2439 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGTGCT 2489

RESULT 14

ADV34986/c

ID ADV34986 standard; cDNA; 185548 BP.

XX AC ADV34986;

XX DT 10-FEB-2005 (first entry)

XX Murine cDNA differentially expressed in the presence of valproate Seq62.
 KW murine; mouse; valproate; ss; multi-parameter high throughput screening;
 KW MPHS; disease signature; neuropsychiatric; neurodegenerative;
 KW schizophrenia; bipolar affective disorder; BAD; autism; Parkinson's;
 KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressa.

XX OS Mus musculus.

XX PN US2003096264-A1.

XX PD 22-MAY-2003.

XX PF 18-JUN-2002; 2002US-00175523.

XX PR 18-JUN-2001; 2001US-0299151P.

XX PR 07-SEP-2001; 2001US-0317828P.

XX PR 25-SEP-2001; 2001US-0325150P.

XX PR 14-NOV-2001; 2001US-0333047P.

XX PR 18-JAN-2002; 2002US-0349936P.

XX PR 04-MAR-2002; 2002US-0361834P.

XX PA (PSYC-) PSYCHIATRIC GENOMICS INC.

XX PI Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;

XX PI Palfreyman M, Rajan P;

XX WIPI; 2004-118903/12.

XX Identifying a compound that can treat disease or disorders, such as, a

PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises

PT determining the expression of one or more efficacy genes in a cell

XX contacted with the test compound.

PS Claim 9; SEQ ID NO 62; 39pp; English.

CC This invention relates to a novel screening method identified as a multi-
CC parameter high throughput screening (MPHTS) assay. Specifically, it
CC refers to an assay that utilizes the disease signature of a plurality of
CC specific genes associated with a particular disease, and identifies
CC differential expression between those cells taken from individuals
CC affected by that disease and those that are not affected. The present
CC invention then describes the screening of candidate pharmaceutical
CC compounds to identify those that have a potential therapeutic benefit for
CC the treatment of neuropsychiatric and neurodegenerative disorders
CC including schizophrenia, bipolar affective disorder (BAD) and autism, as
CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
CC of this invention exhibit various activities including neuroleptic,
CC neurotropic, antimanic and antidepressant. Furthermore, the screening
CC method used in MPHTS will be automated, such that a large number of test
CC compounds may be rapidly screened with a minimal amount of labour and
CC effort. This polynucleotide is the cDNA sequence of a gene that is
CC differentially expressed in mice in the presence of the therapeutic
CC compound valproate, given in an exemplification of the invention.
XX

SQ Sequence 185548 BP; 58272 A; 40343 C; 37771 G; 49162 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.71e+04	Length:	185548
Score:	52.00	Matches:	9
Percent Similarity:	92.9%	Conservative:	4
Best Local Similarity:	64.3%	Mismatches:	1
Query Match:	64.2%	Indels:	0
DB:	13	Gaps:	0

US-10-664-025-3903_COPY_1_17 (1-17) x ADV34986 (1-185548)

OY 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

Db 178598 CARGTACTTTTCTAATATCTGTGTATATGACTGGGGCT 178557

RESULT 15

ADO34435_6/c

Continuation (7 of 7) of ADO34435 from base 600001 (Human SLIT-3 genomic sequence.)
WP Sequence split into 7 fragments LOCUS ADO34435 Accession ADO34435

WP	Fragment Name	Begin	End
WP	ADO34435_0	1	110000
WP	ADO34435_1	100001	210000
WP	ADO34435_2	200001	310000
WP	ADO34435_3	300001	410000
WP	ADO34435_4	400001	510000
WP	ADO34435_5	500001	610000
WP	ADO34435_6	600001	634886

Alignment Scores:

Pred. No.:	7.52e+03	Length:	34886
Score:	49.00	Matches:	8
Percent Similarity:	82.4%	Conservative:	6
Best Local Similarity:	47.1%	Mismatches:	3
Query Match:	60.5%	Indels:	0
DB:	12	Gaps:	0

US-10-664-025-3903_COPY_1_17 (1-17) x ADO34435_6 (1-34886)

OY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

Db 11452 ATGCTAGTATTGTGATCTTTCTTTGGCTATTGTATTATGACAGGTAGT 11402

Search completed: May 9, 2006, 13:30:11
Job time : 92.4 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 13:33:17 ; Search time 22.3 Seconds
(without alignments)
2032.638 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17
Perfect score: 81
Sequence: 1 MALEVLMLLAVLIWTGA 17

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DBV=xlp
-Q=/abs/ABSSWEB_spool/US10664025/runat_08052006_173649_29466/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs03p -USER=US10664025 @CGN 1 1 359 @runat 08052006_173649_29466
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:**
1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/6H COMB.seq:*
6: /cgn2_6/prodata/1/ina/6CTUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length DB	ID	Description
1	81	100.0	469	3	US-09-621-976-1363 Sequence 1363, Ap
2	81	100.0	512	3	US-09-621-976-43 Sequence 43, Appl
3	81	100.0	560	3	US-09-621-976-2587 Sequence 2587, Ap
4	81	100.0	642	3	US-09-621-976-1362 Sequence 1362, Ap
5	81	100.0	1492	3	US-10-104-047-1704 Sequence 1704, Ap
6	54	66.7	372	3	US-09-621-976-15700 Sequence 15700, A
7	47	58.0	564	3	US-09-621-976-25675 Sequence 25675, A
8	47	58.0	1133	3	US-09-270-767-10302 Sequence 10302, A
9	45	55.6	492	3	US-09-252-991A-4051 Sequence 4051, Ap

C 10	45	55.6	601	3	US-09-949-016-163016 Sequence 163016,
C 11	45	55.6	601	3	US-09-949-016-163017 Sequence 163017,
C 12	45	55.6	682	3	US-09-533-559-6972 Sequence 6972, Ap
C 13	45	55.6	1077	3	US-09-252-991A-3983 Sequence 3983, Ap
C 14	45	55.6	1278	3	US-09-252-991A-3896 Sequence 3896, Ap
C 15	45	55.6	1977	3	US-09-252-991A-3924 Sequence 3924, Ap
C 16	45	55.6	2091	3	US-09-252-991A-4016 Sequence 4016, Ap
C 17	45	55.6	2473	3	US-09-949-016-4555 Sequence 4555, Ap
C 18	45	55.6	32211	3	US-09-949-016-16307 Sequence 16307, A
C 19	44	54.3	261	3	US-09-184-418C-106 Sequence 106, App
C 20	44	54.3	261	3	US-10-290-579A-106 Sequence 106, App
C 21	44	54.3	360	3	US-09-513-999C-10116 Sequence 10116, A
C 22	44	54.3	520	3	US-09-621-976-1488 Sequence 1488, Ap
C 23	44	54.3	543	3	US-09-328-352-1729 Sequence 1729, Ap
C 24	44	54.3	1176	3	US-09-583-110-1922 Sequence 1922, Ap
C 25	44	54.3	1200	3	US-09-107-433-1503 Sequence 1503, Ap
C 26	44	54.3	1350	3	US-09-149-476-248 Sequence 248, App
C 27	44	54.3	1521	2	US-09-004-502-2 Sequence 2, Appli
C 28	44	54.3	1521	3	US-09-360-125-2 Sequence 2, Appli
C 29	44	54.3	2633	3	US-09-184-418C-104 Sequence 104, App
C 30	44	54.3	2633	3	US-10-290-579A-104 Sequence 104, App
C 31	44	54.3	2651	3	US-09-184-418C-105 Sequence 105, App
C 32	44	54.3	2651	3	US-10-290-579A-105 Sequence 105, App
C 33	44	54.3	8959	3	US-09-184-418C-11 Sequence 11, Appl
C 34	44	54.3	8959	3	US-10-290-579A-11 Sequence 11, Appl
C 35	43	53.1	261	3	US-09-184-418C-79 Sequence 79, Appl
C 36	43	53.1	261	3	US-10-290-579A-79 Sequence 79, Appl
C 37	43	53.1	825	3	US-09-134-000C-1205 Sequence 1205, Ap
C 38	43	53.1	1113	2	US-08-231-342-5 Sequence 5, Appli
C 39	43	53.1	1419	3	US-09-489-039A-6847 Sequence 6847, Ap
C 40	43	53.1	1644	3	US-09-489-039A-6672 Sequence 6672, Ap
C 41	43	53.1	1762	2	US-08-105-483-300 Sequence 300, App
C 42	43	53.1	1762	2	US-08-709-209-300 Sequence 300, App
C 43	43	53.1	1762	2	US-08-458-101-300 Sequence 300, App
C 44	43	53.1	2383	3	US-10-104-047-995 Sequence 995, App
C 45	43	53.1	2565	3	US-09-949-016-5492 Sequence 5492, Ap

ALIGNMENTS

RESULT 1
US-09-621-976-1363 Application US/09621976
; Sequence 1363, Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1363
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.6999998092651
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN

US-09-621-976-1363
Alignment Scores:
Pred. No.: 3.11e-05 Length: 469
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-10-664-025-3903_COPY_1_17 (1-17) x US-09-621-976-1363 (1-469)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrThrGlyAla 17
DB 137 ATGGCGTTAGAAAGTCTTGATGCTCCTCGTGTCTTGAATTGGACCGGTGCT 187

RESULT 2

US-09-621-976-43
; Sequence 43, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 43
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..511
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-621-976-43

Alignment Scores:

Pred. No.: 3.45e-05 Length: 512
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-621-976-43 (1-512)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrThrGlyAla 17
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGTGTCTTGAATTGGACCGGTGCT 103

RESULT 3

US-09-621-976-2587
; Sequence 2587, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2587
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 372..539
US-09-621-976-2587

Alignment Scores:

Pred. No.: 3.83e-05 Length: 560
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-621-976-2587 (1-560)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrThrGlyAla 17
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGTGTCTTGAATTGGACCGGTGCT 103

RESULT 4

US-09-621-976-1362
; Sequence 1362, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1362
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..526
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-621-976-1362

Alignment Scores:

Pred. No.: 4.49e-05 Length: 642
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-621-976-1362 (1-642)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrThrGlyAla 17
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGTGTCTTGAATTGGACCGGTGCT 103

RESULT 5

US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HL-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492

[illegible]

```
Score: 45.00 Matches: 10
Percent Similarity: 73.3% Conservative: 1
Best Local Similarity: 66.7% Mismatches: 4
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-252-991A-4051 (1-492)
Qy 2 AlaLeuGlValLeuMetLeuLeuAlaValLeuLeuThrGly 16
Db 371 GCCCTGCTGGCCCTGGTGTCTGTGGCGGTGCTGTGGCTGGCGGC 327

RESULT 10
US-09-949-016-163016/c
; Sequence 163016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163016
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-163016

Alignment Scores:
Pred. No.: 65.3 Length: 601
Score: 45.00 Matches: 7
Percent Similarity: 100.0% Conservative: 4
Best Local Similarity: 63.6% Mismatches: 0
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-949-016-163016 (1-601)
Qy 5 ValLeuMetLeuAlaValLeuLeuThr 15
Db 574 ATCCGTGTGCTTGTGCTCTTCTTGTCTGGAGG 542

RESULT 11
US-09-949-016-163017/c
; Sequence 163017, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163017
; LENGTH: 601
; TYPE: DNA
```

```
; ORGANISM: Human
US-09-949-016-163017

Alignment Scores:
Pred. No.: 65.3 Length: 601
Score: 45.00 Matches: 7
Percent Similarity: 100.0% Conservative: 4
Best Local Similarity: 63.6% Mismatches: 0
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-949-016-163017 (1-601)
Qy 5 ValLeuMetLeuAlaValLeuLeuThr 15
Db 206 ATCCGTGTGCTTGTGCTCTTCTTGTCTGGAGG 174

RESULT 12
US-09-533-559-6972
; Sequence 6972, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Key
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6972
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(682)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-6972

Alignment Scores:
Pred. No.: 75.7 Length: 682
Score: 45.00 Matches: 8
Percent Similarity: 75.0% Conservative: 4
Best Local Similarity: 50.0% Mismatches: 4
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-533-559-6972 (1-682)
Qy 1 MetaLaLeuGlValLeuMetLeuLeuAlaValLeuLeuThrGly 16
Db 599 TTGGCGTACATCCATGATCTATTCATGCTCTTGTGAGCGGT 646

RESULT 13
US-09-252-991A-3983/c
; Sequence 3983, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```


; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3983
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (504),(604),(642),(753)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-3983

Alignment Scores:
Pred. No.: 129 Length: 1077
Score: 45.00 Matches: 10
Percent Similarity: 73.3% Conservative: 1
Best Local Similarity: 66.7% Mismatches: 4
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-252-991A-3983 (1-1077)
QY 2 AlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrrThrGly 16
Db 241 GCCCTGCTGCCCTGGTGTCTGCTGGCGGTGCTGGCTGGCTGGGC 197

RESULT 14
US-09-252-991A-3896
; Sequence 3896, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3896
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (12),(50),(150)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-3896

Alignment Scores:
Pred. No.: 158 Length: 1278
Score: 45.00 Matches: 10
Percent Similarity: 73.3% Conservative: 1
Best Local Similarity: 66.7% Mismatches: 4
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-252-991A-3896 (1-1278)
QY 2 AlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrrThrGly 16
Db 413 GCCCTGCTGCCCTGGTGTCTGCTGGCGGTGCTGGCTGGCTGGGC 457

RESULT 15
US-09-252-991A-3924
; Sequence 3924, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3924
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (44),(82),(182)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-3924

Alignment Scores:
Pred. No.: 263 Length: 1977
Score: 45.00 Matches: 10
Percent Similarity: 73.3% Conservative: 1
Best Local Similarity: 66.7% Mismatches: 4
Query Match: 55.6% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-252-991A-3924 (1-1977)
QY 2 AlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrrThrGly 16
Db 445 GCCCTGCTGCCCTGGTGTCTGCTGGCGGTGCTGGCTGGCTGGGC 489

Search completed: May 9, 2006, 14:54:02
Job time : 23.3 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 14:40:21 ; Search time 112.5 Seconds

(without alignments)
1874.392 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17

Perfect score: 81

Sequence: 1 MALEVLMLAVLTGCA 17

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlp
-Q/abes/ABSSWEB spool/US10664025/runat_08052006_173653_29528/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06p
-USER=US10664025 @CGN 1 1 2064 @runat_08052006_173653_29528 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA_Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	470	3	US-09-731-872-24
2	81	100.0	470	3	US-09-876-997-24
3	81	100.0	470	9	US-10-643-836-24
4	81	100.0	691	3	US-09-731-872-19
5	81	100.0	691	3	US-09-876-997-19
6	81	100.0	691	9	US-10-643-836-19
7	81	100.0	1337	8	US-10-467-046-2

8	81	100.0	1480	7	US-10-467-535-16	Sequence 16, Appl
9	81	100.0	1492	6	US-10-104-047-1704	Sequence 1704, Ap
10	81	100.0	1898	5	US-10-050-704-27	Sequence 27, Appl
11	81	100.0	1898	7	US-10-798-512-27	Sequence 27, Appl
12	81	100.0	1898	9	US-10-472-533-175	Sequence 175, App
13	81	100.0	10115	8	US-10-467-046-1	Sequence 1, Appli
14	52	64.2	185548	5	US-10-175-523-62	Sequence 62, Appl
15	52	64.2	185548	10	US-11-099-286-62	Sequence 62, Appl
16	50	61.7	369	8	US-10-425-115-66920	Sequence 66920, A
17	48	59.3	600	9	US-10-972-079-63973	Sequence 63973, A
18	48	59.3	1414	7	US-10-767-701-13037	Sequence 13037, A
19	47	58.0	696	8	US-10-767-795-3017	Sequence 3017, Ap
20	47	58.0	3407	10	US-11-097-143-24373	Sequence 24373, A
21	46	56.8	354	7	US-10-424-598-72	Sequence 72, Appl
22	46	56.8	737	8	US-10-363-345A-6527	Sequence 6527, Ap
23	46	56.8	737	8	US-10-363-345A-6528	Sequence 6528, Ap
24	46	56.8	737	9	US-10-363-483A-6527	Sequence 6527, Ap
25	46	56.8	737	9	US-10-363-483A-6528	Sequence 6528, Ap
26	46	56.8	1003	5	US-10-027-632-121354	Sequence 121354,
27	46	56.8	1003	6	US-10-027-632-121354	Sequence 121354,
28	46	56.8	1642	4	US-09-925-065A-77448	Sequence 77448, A
29	46	56.8	1642	4	US-09-925-065A-77449	Sequence 77449, A
30	46	56.8	1642	4	US-09-925-065A-77450	Sequence 77450, A
31	46	56.8	2082	6	US-10-156-761-4225	Sequence 4225, Ap
32	46	56.8	2459	10	US-11-097-143-23666	Sequence 23666, A
33	46	56.8	4473	8	US-10-425-115-92947	Sequence 92947, A
34	46	56.8	8157	8	US-10-425-115-14619	Sequence 14619, A
35	46	56.8	20460	10	US-11-097-143-23665	Sequence 23665, A
36	46	56.8	20460	10	US-11-097-143-28396	Sequence 28396, A
37	46	56.8	2256646	7	US-10-470-565-1	Sequence 1, Appli
38	46	56.8	9025608	6	US-10-156-761-1	Sequence 1, Appli
39	45	55.6	243	3	US-09-998-598-133	Sequence 133, App
40	45	55.6	499	5	US-10-027-632-137049	Sequence 137049,
41	45	55.6	499	5	US-10-027-632-137050	Sequence 137050,
42	45	55.6	499	6	US-10-027-632-137049	Sequence 137049,
43	45	55.6	499	6	US-10-027-632-137050	Sequence 137050,
44	45	55.6	500	6	US-10-029-386-9086	Sequence 9086, Ap
45	45	55.6	507	7	US-10-021-323-12898	Sequence 12898, A

ALIGNMENTS

RESULT 1

US-09-731-872-24
; Sequence 24, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueloret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLMLAVLTGCA/EN
US-09-731-872-24

Alignment Scores:
Pred. No.: 0.000108 Length: 470
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-731-872-24 (1-470)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrGlyAla 17
Db 137 ATGGCGTTAGAGTCTTGATGCTCTCGTCTTGTATTGGACCGGTGCT 187

RESULT 2
; Sequence 24, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR FILING DATE: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-876-997-24

Alignment Scores:
Pred. No.: 0.000108 Length: 470
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-876-997-24 (1-470)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrGlyAla 17
Db 137 ATGGCGTTAGAGTCTTGATGCTCTCGTCTTGTATTGGACCGGTGCT 187

RESULT 3
; Sequence 24, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG

; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-10-643-836-24

Alignment Scores:
Pred. No.: 0.000108 Length: 470
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-643-836-24 (1-470)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrGlyAla 17
Db 137 ATGGCGTTAGAGTCTTGATGCTCTCGTCTTGTATTGGACCGGTGCT 187

RESULT 4
; Sequence 19, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-731-872-19

Alignment Scores:
Pred. No.: 0.00017 Length: 691

Score: 81.00 Matches: 17
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 3
Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-876-997-19 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrGlyAla 17
DB 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCT 92

RESULT 5
US-09-876-997-19
; Sequence 19, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pbm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-876-997-19

Alignment Scores:
Pred. No.: 0.00017 Length: 691
Score: 81.00 Matches: 17
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 3
Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-876-997-19 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrGlyAla 17
DB 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCT 92

RESULT 6
US-10-643-836-19
; Sequence 19, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872

; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pbm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-10-643-836-19

Alignment Scores:
Pred. No.: 0.00017 Length: 691
Score: 81.00 Matches: 17
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 3
Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-643-836-19 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuThrGlyAla 17
DB 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCT 92

RESULT 7
US-10-467-046-2
; Sequence 2, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (53)..(103)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (104)..(526)
US-10-467-046-2

Alignment Scores:
Pred. No.: 0.000371 Length: 1337
Score: 81.00 Matches: 17
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%

```
DB: 8 Gaps: 0
US-10-664-025-3903_COPY_1_17 (1-17) x US-10-467-046-2 (1-1337)
Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 53 ATGGCGTTAGAACTCTCGTCTGCTTGTCTTGTGACCGGTGCT 103

RESULT 8
US-10-467-535-16
; Sequence 16, Application US/10467535
; Publication No. US20040146970A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; YAO, Monique G.
; APPLICANT: ISON, Craig H.; LU, Yan
; APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
; APPLICANT: BAUGHN, Mariah R.; DING, Li
; APPLICANT: XU, Yuming; GRETZEN, Kimberly J.
; APPLICANT: TANG, Tom Y.; LAL, Preeti G.
; APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
; APPLICANT: LU, Dyung Aina M.; RICHARDSON, Thomas W.
; APPLICANT: TRAN, Uyen K.; KHARE, Reena
; APPLICANT: CHAWLA, Narinder K.
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-0903 USN
; CURRENT APPLICATION NUMBER: US/10/467,535
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/US02/037115
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/268,111
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/271,175
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,503
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/274,552
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 1480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5284076CB1
US-10-467-535-16

Alignment Scores:
Pred. No.: 0.000419 Length: 1480
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-467-535-16 (1-1480)
Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 29 ATGGCGTTAGAACTCTCGTCTGCTTGTCTTGTGACCGGTGCT 79

RESULT 9
US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:

; PRIOR FILING DATE:
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1704

Alignment Scores:
Pred. No.: 0.000423 Length: 1492
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-104-047-1704 (1-1492)
Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 52 ATGGCGTTAGAACTCTCGTCTGCTTGTCTTGTGACCGGTGCT 102

RESULT 10
US-10-050-704-27
; Sequence 27, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-050-704-27

Alignment Scores:
Pred. No.: 0.000561 Length: 1898
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-050-704-27 (1-1898)
Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 15 ATGGCGTTAGAACTCTCGTCTGCTTGTCTTGTGACCGGTGCT 65

RESULT 11
US-10-798-512-27
; Sequence 27, Application US/10798512
```

```
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-798-512-27

Alignment Scores:
Pred. No.: 0.000561 Length: 1898
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-798-512-27 (1-1898)

Qy 1 MetaLeuGlValLeuMetLeuLeuAlaValLeuIleThrGlyVala 17
Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGTGTCTTGATTGACCGGTGCT 65

RESULT 12
US-10-472-533-175
; Sequence 175, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS906PCT
; CURRENT APPLICATION NUMBER: US/10/472,533
; CURRENT FILING DATE: 2003-09-20
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 650
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1398)..(1398)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature

US-10-472-533-175
; LOCATION: (1428)..(1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-472-533-175

Alignment Scores:
Pred. No.: 0.000561 Length: 1898
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-472-533-175 (1-1898)

Qy 1 MetaLeuGlValLeuMetLeuLeuAlaValLeuIleThrGlyVala 17
Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGTGTCTTGATTGACCGGTGCT 65

RESULT 13
US-10-467-046-1
; Sequence 1, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Gid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 10115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2387...2501
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5443...5646
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6649...6747
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8907...9774
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2986)..(2986)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3847)..(3848)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3879)..(3879)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4105)..(4105)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4107)..(4109)
; OTHER INFORMATION: n is a, c, g, or t
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7347)..(7348)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8197)..(8197)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8293)..(8293)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8334)..(8336)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9942)..(9942)
; OTHER INFORMATION: n is a, c, g, or t
US-10-467-046-1

Alignment Scores:
Pred. No.: 0.00404 Length: 10115
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-467-046-1 (1-10115)

Qy 1 MetalaLeuGluValLeuMetLeuAlaValLeuIleThrGlyAla 17
Db 2439 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGTTGATTGGACCGGTGCT 2489

RESULT 14
US-10-175-523-62/c
; Sequence 62, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/1J795-US3
; CURRENT FILING DATE: 2002-06-18
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/361,834
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 185548
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-099-266-62

Alignment Scores:
Pred. No.: 1.72e+04 Length: 185548
Score: 52.00 Matches: 9
Percent Similarity: 92.9% Conservative: 4
Best Local Similarity: 64.3% Mismatches: 1
Query Match: 64.2% Indels: 0
DB: 10 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-11-099-266-62 (1-185548)

Qy 4 GluValLeuMetLeuAlaValLeuIleThrGlyAla 17
Db 178598 CAAGTACTTTTCTAATATCTGTTGTTATATGACTGGGCT 178557

Search completed: May 9, 2006, 15:22:07
Job time : 126.5 secs
```

```
; ORGANISM: Mus musculus
US-10-175-523-62

Alignment Scores:
Pred. No.: 1.72e+04 Length: 185548
Score: 52.00 Matches: 9
Percent Similarity: 92.9% Conservative: 4
Best Local Similarity: 64.3% Mismatches: 1
Query Match: 64.2% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-175-523-62 (1-185548)

Qy 4 GluValLeuMetLeuAlaValLeuIleThrGlyAla 17
Db 178598 CAAGTACTTTTCTAATATCTGTTGTTATATGACTGGGCT 178557

RESULT 15
US-11-099-266-62/c
; Sequence 62, Application US/11099266
; Publication No. US2005018143A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 03235/100J795-US4
; CURRENT APPLICATION NUMBER: US/11/099,266
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: US 10/175,523
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 185548
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-099-266-62

Alignment Scores:
Pred. No.: 1.72e+04 Length: 185548
Score: 52.00 Matches: 9
Percent Similarity: 92.9% Conservative: 4
Best Local Similarity: 64.3% Mismatches: 1
Query Match: 64.2% Indels: 0
DB: 10 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-11-099-266-62 (1-185548)

Qy 4 GluValLeuMetLeuAlaValLeuIleThrGlyAla 17
Db 178598 CAAGTACTTTTCTAATATCTGTTGTTATATGACTGGGCT 178557

Search completed: May 9, 2006, 15:22:07
Job time : 126.5 secs
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: May 9, 2006, 14:45:03 ; Search time 75.8 Seconds
(without alignments)
1370.049 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17
Perfect score: 81
Sequence: 1 MALEVLMLLAVLTGTA 17

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+pn.model -DEV=xlh
-Q=/abss/ABSSWEB.spool/US10664025/runat.08052006.173656.29587/app.query.fasta.1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss808
-USER=US10664025 @CGN 1 1 981 @runat.08052006.173656.29587 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : Published Applications NA New:

- 1: /SIDSS/ptodata/1/pubpna/US08_NEW_PUB.seq1.*
- 2: /SIDSS/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /SIDSS/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /SIDSS/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /SIDSS/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 6: /SIDSS/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 7: /SIDSS/ptodata/1/pubpna/US09_NEW_PUB.seq1.*
- 8: /SIDSS/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
- 9: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 10: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq1.*
- 11: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq2.*
- 12: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq3.*
- 13: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq4.*
- 14: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 15: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
- 16: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
- 17: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq4.*
- 18: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq5.*
- 19: /SIDSS/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	81	100.0	497	11	US-10-475-075-564	Sequence 564, App
2	81	100.0	500	11	US-10-475-075-563	Sequence 563, App
3	81	100.0	1492	18	US-11-072-512-1704	Sequence 1704, Ap
C 4	46	56.8	1642	7	US-09-925-065A-77448	Sequence 77448, A
C 5	46	56.8	1642	7	US-09-925-065A-77449	Sequence 77449, A
C 6	46	56.8	1642	7	US-09-925-065A-77450	Sequence 77450, A
C 7	46	56.8	1642	11	US-10-301-480-178687	Sequence 178687,
C 8	46	56.8	1642	11	US-10-301-480-178688	Sequence 178688,
C 9	46	56.8	1642	11	US-10-301-480-178689	Sequence 178689,
C 10	46	56.8	1642	12	US-10-301-480-792096	Sequence 792096,
C 11	46	56.8	1642	12	US-10-301-480-792097	Sequence 792097,
C 12	46	56.8	1642	12	US-10-301-480-792098	Sequence 792098,
C 13	45	55.6	294	11	US-10-932-182A-166510	Sequence 166510,
C 14	45	55.6	294	11	US-10-932-182A-166510	Sequence 166510,
C 15	45	55.6	492	17	US-11-098-686-9464	Sequence 9464, Ap
C 16	45	55.6	563	7	US-09-925-065A-122833	Sequence 122833,
C 17	45	55.6	563	7	US-09-925-065A-122834	Sequence 122834,
C 18	45	55.6	569	7	US-09-925-065A-857547	Sequence 857547,
C 19	45	55.6	570	7	US-09-925-065A-857555	Sequence 857555,
C 20	45	55.6	585	12	US-10-301-480-220835	Sequence 220835,
C 21	45	55.6	585	12	US-10-301-480-834244	Sequence 834244,
C 22	45	55.6	585	12	US-10-301-480-834244	Sequence 834244,
C 23	45	55.6	585	12	US-10-301-480-834245	Sequence 834245,
C 24	45	55.6	587	7	US-09-925-065A-432235	Sequence 432235,
C 25	45	55.6	587	7	US-09-925-065A-432236	Sequence 432236,
C 26	45	55.6	588	12	US-10-301-480-494737	Sequence 494737,
C 27	45	55.6	588	12	US-10-301-480-494738	Sequence 494738,
C 28	45	55.6	588	12	US-10-301-480-1108146	Sequence 1108146,
C 29	45	55.6	588	12	US-10-301-480-1108147	Sequence 1108147,
C 30	45	55.6	628	12	US-10-301-480-537863	Sequence 537863,
C 31	45	55.6	628	12	US-10-301-480-1151272	Sequence 1151272,
C 32	45	55.6	630	7	US-09-925-065A-845855	Sequence 845855,
C 33	45	55.6	632	7	US-09-925-065A-848053	Sequence 848053,
C 34	45	55.6	946	7	US-09-925-065A-725112	Sequence 725112,
C 35	45	55.6	193084	17	US-11-121-086-82	Sequence 82, Appl
C 36	45	55.6	1457619	17	US-11-098-686-8739	Sequence 8739, Ap
C 37	44	54.3	532	7	US-09-925-065A-67585	Sequence 67585, A
C 38	44	54.3	532	11	US-10-301-480-168824	Sequence 168824,
C 39	44	54.3	532	12	US-10-301-480-782233	Sequence 782233,
C 40	44	54.3	561	12	US-10-301-480-285522	Sequence 285522,
C 41	44	54.3	561	12	US-10-301-480-285523	Sequence 285523,
C 42	44	54.3	561	12	US-10-301-480-285525	Sequence 285525,
C 43	44	54.3	561	12	US-10-301-480-898931	Sequence 898931,
C 44	44	54.3	561	12	US-10-301-480-898932	Sequence 898932,
C 45	44	54.3	561	12	US-10-301-480-898934	Sequence 898934,

ALIGNMENTS

RESULT 1
US-10-475-075-564
; Sequence 564, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 564
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: CDS
; LOCATION: 52..495
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 52..102
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-10-475-075-564

Alignment Scores:
Pred. No.: 2.98e-05 Length: 497
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-475-075-564 (1-497)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 52 ATGGCGTTAGAGTCTTGATGCTCTCGTCTCTTGATTGGACCGGTGCT 102

RESULT 2

US-10-475-075-563
; Sequence 563, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dunas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 563
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 72..500
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 72..122
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1..3
; OTHER INFORMATION: n = a, g, c or t
US-10-475-075-563

Alignment Scores:
Pred. No.: 3e-05 Length: 500
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-475-075-563 (1-500)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

Db 72 ATGGCGTTAGAGTCTTGATGCTCTCGTCTCTTGATTGGACCGGTGCT 122

RESULT 3

US-11-072-512-1704
; Sequence 1704, Application US/11072512
; Publication No. US2006002945A1
; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEXI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-1704

Alignment Scores:

Pred. No.: 0.000107 Length: 1492
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 18 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-11-072-512-1704 (1-1492)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 52 ATGGCGTTAGAGTCTTGATGCTCTCGTCTCTTGATTGGACCGGTGCT 102

RESULT 4

US-09-925-065A-77448/c
; Sequence 77448, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

```
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77448
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-77448

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-925-065A-77448 (1-1642)
Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 1160 AAAGTTCCTCGTGTGTCGTCTCCCTGGACTGGGGCC 1119

RESULT 5
US-09-925-065A-77449/c
; Sequence 77449, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77449
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-77449

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-925-065A-77449 (1-1642)
Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 1160 AAAGTTCCTCGTGTGTCGTCTCCCTGGACTGGGGCC 1119

RESULT 6
US-09-925-065A-77450/c
; Sequence 77450, Application US/09925065A
; Publication No. US20040181048A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77450
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-77450

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-09-925-065A-77450 (1-1642)
Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
Db 1160 AAAGTTCCTCGTGTGTCGTCTCCCTGGACTGGGGCC 1119

RESULT 7
US-10-301-480-178687/c
; Sequence 178687, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178687
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-178687

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-301-480-178687 (1-1642)
Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
```

Db 1160 AAGTTCTCCTGTTGTTGCTGCTCCCTGGACTGGGGCC 1119
:::|||||:::|||||:::|||||

RESULT 8
US-10-301-480-178688/c
; Sequence 178688, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178688
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-178688

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-301-480-178688 (1-1642)

Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
:::|||||:::|||||:::|||||

Db 1160 AAGTTCTCCTGTTGTTGCTGCTCCCTGGACTGGGGCC 1119
:::|||||:::|||||:::|||||

RESULT 9
US-10-301-480-178689/c
; Sequence 178689, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178689
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-178689

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-301-480-178689 (1-1642)

Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
:::|||||:::|||||:::|||||

Db 1160 AAGTTCTCCTGTTGTTGCTGCTCCCTGGACTGGGGCC 1119
:::|||||:::|||||:::|||||

RESULT 10
US-10-301-480-792096/c
; Sequence 792096, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 792096
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-792096

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 12 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-301-480-792096 (1-1642)

Qy 4 GluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17
:::|||||:::|||||:::|||||

Db 1160 AAGTTCTCCTGTTGTTGCTGCTCCCTGGACTGGGGCC 1119
:::|||||:::|||||:::|||||

RESULT 11
US-10-301-480-792097/c
; Sequence 792097, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 792097
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-792097

Alignment Scores:
Pred. No.: 406 Length: 1642
Score: 46.00 Matches: 9
Percent Similarity: 85.7% Conservative: 3
Best Local Similarity: 64.3% Mismatches: 2
Query Match: 56.8% Indels: 0
DB: 12 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x US-10-301-480-792097 (1-1642)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 13:30:21 ; Search time 597.7 Seconds

(without alignments)
1996.103 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17

Perfect score: 81

Sequence: 1 MALEVLMLLAVLIWTGA 17

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/abses/ABSSWEB spool/US1064025/runat_08052006_173646_29404/app query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abses07
-USER=US1064025 @CGN 1.1 6731 @runat_08052006_173646_29404 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	81	100.0	331	1 A1138881	A1138881 qd99c04.x
2	81	100.0	516	2 BE972717	BE972717 601652136
3	81	100.0	626	3 B1463012	B1463012 603204535
C 4	81	100.0	753	5 BU561351	BU561351 AGENCOURT
5	81	100.0	776	2 BG772527	BG772527 602720619
6	81	100.0	778	3 B1459538	B1459538 603200545
7	81	100.0	878	3 B1520154	B1520154 603071295

8	62	76.5	598	2	BG081002	BG081002 H3059H12-
9	62	76.5	629	5	BY731676	BY731676 H3059H12-
10	62	76.5	862	7	CO810251	CO810251 AGENCOURT
11	59	72.8	577	5	BX118111	BX118111 AGENCOURT
12	54	66.7	589	6	CA289989	CA289989 SCAGFL801
13	54	66.7	1009	6	CA278912	CA278912 SCBFRT309
14	53	65.4	863	2	BG741530	BG741530 602635139
15	52	64.2	595	1	AW011198	AW011198 STL7H06 P
16	52	64.2	1056	8	DR154376	DR154376 49153495
17	52	64.2	1084	8	DR139602	DR139602 49153495
18	52	64.2	1162	8	DR154420	DR154420 49139733
19	51	63.0	689	9	CE251157	CE251157 tigr-q88-
C 20	50	61.7	719	10	CNS01EXN	AL141132 Anopheles
21	50	61.7	866	2	BF571847	BF571847 602076741
22	49	60.5	556	9	AZ624138	AZ624138 IM0462M06
23	49	60.5	721	10	CG350468	CG350468 CGSANO6TC
24	49	60.5	867	6	CF715063	CF715063 CCADV37TF
25	49	60.5	915	10	CG212053	CG212053 OGZBC45TV
26	48	59.3	260	9	AZ558974	AZ558974 RPCI-23-1
27	48	59.3	301	1	AV013267	AV013267 AV013267
28	48	59.3	341	1	AW428893	AW428893 Ljlrnpeset
29	48	59.3	474	1	AV422253	AV422253 AV422253
30	48	59.3	503	2	BI099005	BI099005 IPI_36 B1
31	48	59.3	504	2	BG102437	BG102437 RHIZ2_23
32	48	59.3	514	6	CB829260	CB829260 LJNEST96a
33	48	59.3	535	6	CA121482	CA121482 SCEZLR100
34	48	59.3	537	1	AW719616	AW719616 LJNEST6h7
35	48	59.3	552	3	BM324073	BM324073 PIC1_24 H
36	48	59.3	557	6	CB828146	CB828146 LJNEST83C
37	48	59.3	579	6	CA208211	CA208211 SCWCSB111
38	48	59.3	580	2	BE594345	BE594345 P11_32 C0
39	48	59.3	596	6	CD431906	CD431906 ETH1_11 B
C 40	48	59.3	620	1	AJ414707	AJ414707 AJ414707
41	48	59.3	623	9	BH392706	BH392706 AG-ND-102
42	48	59.3	658	6	CA143424	CA143424 SCQSR203
43	48	59.3	659	6	CA226596	CA226596 SCRTFL300
44	48	59.3	672	6	CA290401	CA290401 SCUTSD102
45	48	59.3	703	1	AW720591	AW720591 LJNEST16b

ALIGNMENTS

RESULT 1

A1138881/c

LOCUS

DEFINITION

3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 395 Std Error: 0.00

Seq primer: -40m13 fwd, ET from Amersham.

Location/Qualifiers

1. .331

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1737606"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer 15'
TGTTCACCAATCGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.: 0.000432 Length: 331
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x A1138881 (1-331)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
Db 331 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGAATTGGACCGGTGCT 281

RESULT 2

BE972717 516 bp mRNA linear EST 04-OCT-2000
LOCUS 601652136F1 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:3935288 5',
DEFINITION mRNA sequence.

ACCESSION BE972717 GI:10586053

VERSION EST.

KEYWORDS Homo sapiens

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 516)

NIH-MGC http://mhc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: CLONETECH Laboratories, Inc.

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCW777 row: h column: 09

High quality sequence stop: 505.

Location/Qualifiers

1..516

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3935288"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_82"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:

SfiI (ggccgctcgcc); Site 2: SfiI (ggccattatggcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGGCCATTATGCCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGCGCCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."

ORIGIN

Alignment Scores:
Pred. No.: 0.000728 Length: 516
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BE972717 (1-516)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

Db 120 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGAATTGGACCGGTGCT 170

RESULT 3

BE972717

LOCUS 603204535F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5270384 5',

DEFINITION mRNA sequence.

ACCESSION BE972717 GI:15253668

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 626)

NIH-MGC http://mhc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLA111682 row: e column: 09

High quality sequence stop: 624.

Location/Qualifiers

1..626

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5270384"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/note="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTTN-3', size-selected for average

insert size 2.2 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 0.000915 Length: 626
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BU561351 (1-626)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAla 17
|||||
Db 57 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 107

RESULT 4
BU561351/c

LOCUS BU561351 753 bp mRNA linear EST 16-SEP-2002

DEFINITION AGENCOURT_10278669 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6592356

ACCESSION BU561351

VERSION BU561351

KEYWORDS BU561351.1 GI:22911647

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 753)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LCM2815 row: 9 column: 12
High quality sequence stop: 442.

FEATURES
Location/Qualifiers
1..753
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6592356"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggcattggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Alignment Scores:
Pred. No.: 0.00114 Length: 753
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BU561351 (1-753)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAla 17
|||||

Db 316 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 266

RESULT 5
BG772527

LOCUS BG772527 776 bp mRNA linear EST 15-MAY-2001

DEFINITION 602720619F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837612 5', mRNA sequence.

ACCESSION BG772527

VERSION BG772527.1 GI:14083180

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 776)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10770 row: e column: 05
High quality sequence stop: 773.

FEATURES
Location/Qualifiers
1..776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4837612"
/lab_host="DH10B"
/clone_lib="NIH MGC 97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 0.00118 Length: 776
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BG772527 (1-776)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAla 17
|||||

Db 57 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 107

RESULT 6
BI459538

LOCUS BI459538 778 bp mRNA linear EST 21-AUG-2001

DEFINITION 603200545F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5266642 5', mRNA sequence.

ACCESSION BI459538

VERSION BI459538.1 GI:15250194

KEYWORDS EST.

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE   1 (bases 1 to 778)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11672 row: i column: 11
            High quality sequence stop: 674.
FEATURES   Location/Qualifiers
            source
            1..778
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5266642"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_97"
            /note="Organ: testis; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (gtcag); Oligo-dT primed using primer
            5'-TTTTTTTTTTTNN-3', size-selected for average
            insert size 2.2 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      0.00118      Length:      778
Score:          81.00      Matches:      17
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             3          Gaps:         0

US-10-664-025-3903_COPY_1_17 (1-17) x BI459538 (1-778)

QY      1 MetAlaLeuGlValLeuMetLeuAlaValLeuLeuTrrThrGlyAla 17
Db      191 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGTTGATTGACCGTGCT 241

RESULT 7
BIS20154      878 bp      mRNA      linear      EST 29-AUG-2001
LOCUS        BIS20154
DEFINITION   mRNA sequence.
ACCESSION   BIS20154
VERSION     BIS20154.1 GI:15344946
KEYWORDS     EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE   1 (bases 1 to 878)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11672 row: i column: 11
            High quality sequence stop: 674.
FEATURES   Location/Qualifiers
            source
            1..778
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5266642"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_97"
            /note="Organ: testis; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (gtcag); Oligo-dT primed using primer
            5'-TTTTTTTTTTTNN-3', size-selected for average
            insert size 2.2 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      0.00118      Length:      778
Score:          81.00      Matches:      17
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             3          Gaps:         0

US-10-664-025-3903_COPY_1_17 (1-17) x BI459538 (1-778)

QY      1 MetAlaLeuGlValLeuMetLeuAlaValLeuLeuTrrThrGlyAla 17
Db      191 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGTTGATTGACCGTGCT 241

RESULT 7
BIS20154      878 bp      mRNA      linear      EST 29-AUG-2001
LOCUS        BIS20154
DEFINITION   mRNA sequence.
ACCESSION   BIS20154
VERSION     BIS20154.1 GI:15344946
KEYWORDS     EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE   1 (bases 1 to 878)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11672 row: i column: 11
            High quality sequence stop: 674.
FEATURES   Location/Qualifiers
            source
            1..778
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5266642"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_97"
            /note="Organ: testis; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (gtcag); Oligo-dT primed using primer
            5'-TTTTTTTTTTTNN-3', size-selected for average
            insert size 2.2 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      0.00136      Length:      878
Score:          81.00      Matches:      17
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             3          Gaps:         0

US-10-664-025-3903_COPY_1_17 (1-17) x BI520154 (1-878)

QY      1 MetAlaLeuGlValLeuMetLeuAlaValLeuLeuTrrThrGlyAla 17
Db      30 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGTTGATTGACCGTGCT 80

RESULT 8
BG081002      598 bp      mRNA      linear      EST 17-DEC-2003
LOCUS        BG081002
DEFINITION   H3059H12-5 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION   BG081002
VERSION     BG081002.2 GI:40015217
KEYWORDS     EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 598)
AUTHORS    Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Doi,H.,
            Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
            Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
            Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
            Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
            10922068
            On Jan 26, 2001 this sequence version replaced gi:12563570.
            Other ESTs: H3059H12-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@gsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please

```

```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11405 row: m column: 24
High quality sequence stop: 851.
FEATURES   Location/Qualifiers
            source
            1..878
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5163455"
            /tissue_type="medulla"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_119"
            /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: EcoRV (destroyed); RNA source normal medulla from
            anonymous male age 27. Library is oligo-dT primed and
            directionally cloned (EcoRV site is destroyed upon
            cloning). Average insert size 1.3 kb, insert size range
            0.9-3 kb. Library is normalized and enriched for
            full-length clones and was constructed by C. Gruber
            (Invitrogen). Research Genetics tracking code 013. Note:
            this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      0.00136      Length:      878
Score:          81.00      Matches:      17
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             3          Gaps:         0

US-10-664-025-3903_COPY_1_17 (1-17) x BI520154 (1-878)

QY      1 MetAlaLeuGlValLeuMetLeuAlaValLeuLeuTrrThrGlyAla 17
Db      30 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGTTGATTGACCGTGCT 80

RESULT 8
BG081002      598 bp      mRNA      linear      EST 17-DEC-2003
LOCUS        BG081002
DEFINITION   H3059H12-5 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION   BG081002
VERSION     BG081002.2 GI:40015217
KEYWORDS     EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 598)
AUTHORS    Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Doi,H.,
            Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
            Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
            Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
            Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
            10922068
            On Jan 26, 2001 this sequence version replaced gi:12563570.
            Other ESTs: H3059H12-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@gsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please

```

visit <http://lgaun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3059 row: H column: 12
Seq primer: -21M13 Reverse
High quality sequence stop: 598
POLYA=No.

FEATURES

source Location/Qualifiers
1..598
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:H3059H12-5"
/db_xref="taxon:10090"
/clone="H3059H12"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Alignment Scores:
Pred. No.: 1..91 Length: 598
Score: 62.00 Matches: 13
Percent Similarity: 88.2% Conservative: 2
Best Local Similarity: 76.5% Mismatches: 2
Query Match: 76.5% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BG081002 (1-598)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla 17
Db 41 ATGGCTTTAGAGCTCTGGTTTATCTGCTGTCTGTTGGACTTGTGCT 91

RESULT 9
BY731676 629 bp mRNA linear EST 17-DEC-2002
LOCUS BY731676
DEFINITION Ovary Mus musculus cDNA clone E330031A07 5', mRNA sequence.
ACCESSION BY731676
VERSION BY731676.1 GI:27144803
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 629)

REFERENCE
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusnic, V.,

Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyszynski, B., Yang, Y., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12456851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
Location/Qualifiers
1..629
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E330031A07"
/sex="female"
/tissue_type="ovary"
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 2 days pregnant adult female ovary"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia"

Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCAACATCGAGTGTTCCTTTTTCCTTTCVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGATTAAATTAATCCCGCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:
Pred. No.: 2.02 Length: 629
Score: 62.00 Matches: 13
Percent Similarity: 88.2% Conservative: 2
Best Local Similarity: 76.5% Mismatches: 2
Query Match: 76.5% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x BY731676 (1-629)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTTrpThrGlyAla 17
DB 46 ATGGCTTTAGAGTCTGCTGTTATCTTCTGCTGCTTGGTTGGACTTGTGCT 96

RESULT 10
C0810251
LOCUS C0810251 862 bp mRNA linear EST 06-AUG-2004
DEFINITION IMAGE:30258260 NIH_MGC_257 Mus musculus cDNA clone
IMAGE:30931566 5', mRNA sequence.
ACCESSION C0810251
VERSION C0810251.1 GI:51028724
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Kathleen Horner, Stanford University
cDNA Library Preparation: Express Genomics
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAm179 row: h column: 07
High quality sequence stop: 654.
Location/Qualifiers
1. .862
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30931566"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_257"
/note="Organ: oocyte; Vector: pExpress-1; Site: 1: EcoRV; Site 2: NotI; cDNA was primed using oligo-dT primer: 5'-pGACTGATCTAGTCGAGCGCGGCCCTT 25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >0.5 kb resulted in an average insert size of 1.0kb. This is a

FEATURES
source

normalized library (primary library is NIH_MGC_256) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library"

ORIGIN

Alignment Scores:
Pred. No.: 2.93 Length: 862
Score: 62.00 Matches: 13
Percent Similarity: 88.2% Conservative: 2
Best Local Similarity: 76.5% Mismatches: 2
Query Match: 76.5% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903_COPY_1_17 (1-17) x C0810251 (1-862)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTTrpThrGlyAla 17
DB 8 ATGGCTTTAGAGTCTGCTGTTATCTTCTGCTGCTTGGTTGGACTTGTGCT 58

RESULT 11

LOCUS BX118111 577 bp mRNA linear EST 07-FEB-2003
DEFINITION BX118111 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:998D122574
IMAGE:1030619, mRNA sequence.
ACCESSION BX118111
VERSION BX118111.1 GI:27881212
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS Ebert,J., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radelof,U., Schneider,D. and Korn,B.
HUMAN UNIGENESET - RZPD3
UNPUBLISHED (2003)
CONTACT: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE:998D122574.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
<http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi?response?libNo=972> Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGCAACAGCTATGAC.
Location/Qualifiers
1. .577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998D122574 ; IMAGE:1030619"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo (dT) primer [5', TGTTCACCAATCTGAAGTCGAGCGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization to C0t5, and was constructed by Bento Soares and M. Fatima Bonaldo. "

FEATURES
source

normalized library (primary library is NIH_MGC_256) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library"

[illegible]

BG741530
LOCUS 602635139F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780232 5', linear EST 15-MAY-2001
DEFINITION mRNA sequence.
ACCESSION BG741530
VERSION BG741530.1 GI:14052183
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 863)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM10638 row: n column: 09
High quality sequence stop: 863.
FEATURES
source
1..863
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4780232"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Alignment Scores:
Pred. No.: 112 Length: 863
Score: 53.00 Matches: 8
Percent Similarity: 88.2% Conservative: 7
Best Local Similarity: 47.1% Mismatches: 2
Query Match: 65.4% Indels: 0
DB: 2 Gaps: 0
US-10-664-025-3903_COPY_1_17 (1-17) x BG741530 (1-863)
Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuLeuTrpThrGlyAla 17
Db 637 TTGGCAITGGAGGTGATCCTTTTATGGAACAGATTTTATGCAGTGGCTCG 687
RESULT 15
AW011198
LOCUS ST17H06 Pine Triplex shoot tip library Pinus taeda cDNA clone
DEFINITION ST17H06, mRNA sequence.
ACCESSION AW011198
VERSION AW011198.1 GI:5859976
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 595)
REFERENCE Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.
AUTHORS The Pine Gene Discovery Project
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall,
Raleigh, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@unity.ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
FEATURES
source
1..595
Location/Qualifiers
/organism="Pinus taeda"
/mol_type="mRNA"
/db_xref="taxon:3352"
/clone="ST17H06"
/lab_host="E. coli BM25.8"
/clone_lib="Pine Triplex shoot tip library"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site 1:
SfiI (A); Site 2: SfiI (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
ORIGIN
Alignment Scores:
Pred. No.: 109 Length: 595
Score: 52.00 Matches: 9
Percent Similarity: 93.3% Conservative: 5
Best Local Similarity: 60.0% Mismatches: 1
Query Match: 64.2% Indels: 0
DB: 1 Gaps: 0
US-10-664-025-3903_COPY_1_17 (1-17) x AW011198 (1-595)
Qy 3 LeuGluValLeuMetLeuLeuAlaValLeuLeuLeuTrpThrGlyAla 17
Db 512 TTGCAGTTGCTGCACATCTTGGCACTGCTTTGGACAGGTGCA 556
Search completed: May 9, 2006, 15:10:12
Job time : 600.7 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:10:22 ; Search time 11.7 Seconds
(without alignments)
638.414 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17
Perfect score: 81
Sequence: 1 MALEVLMLLAFLIWTGA 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	106	4	AAG89145 Human sec
2	81	100.0	143	6	ABP76237 Human GEN
3	81	100.0	148	6	ABP76238 Human GEN
4	81	100.0	158	4	AAG89140 Human sec
5	81	100.0	158	5	ABG97353 Human CGD
6	81	100.0	158	5	ABG31324 Human GSS
7	81	100.0	158	6	ABR47931 Human sec
8	81	100.0	158	6	ABR00175 Human sec
9	81	100.0	158	7	ADB91675 Human gen
10	81	100.0	158	7	ADB65520 Human pro
11	81	100.0	158	7	ADC74338 Human sec
12	81	100.0	159	3	AAB38337 Human sec
13	46	56.8	80	9	ADK40751 HIV Vpu p
14	45	55.6	80	9	ADK40648 HIV Vpu p
15	45	55.6	81	9	ADK40735 HIV Vpu p
16	45	55.6	104	5	ABB94049 Human sec
17	45	55.6	104	5	ABB94025 Human sec
18	45	55.6	104	5	ABG65110 Human alb
19	45	55.6	104	5	ABG65111 Human alb
20	45	55.6	104	8	ADL78378 Albumin f
21	45	55.6	120	4	AAU20830 Human nov
22	45	55.6	120	4	AAU20830 Human oif
23	45	55.6	335	4	AAG72080 Human oif
24	45	55.6	570	8	ADQ03086 P. aerugi

25	45	55.6	658	7	ABO71749 Pseudomon
26	45	55.6	963	6	ABU33173 Protein e
27	45	55.6	983	9	ABE41559 L. pneumo
28	45	55.6	984	9	ABE38276 L. pneumo
29	44	54.3	33	3	AAB69261 HIV-1 non
30	44	54.3	86	3	AAB69344 HIV-1 non
31	44	54.3	134	9	ADZ88786 Breast sp
32	44	54.3	180	6	ADA34568 Acinetoba
33	44	54.3	306	6	ABR58443 Human NOV
34	44	54.3	321	6	ABR58442 Human NOV
35	44	54.3	321	6	ABR58445 Human NOV
36	44	54.3	321	8	ADJ57406 Secretary
37	44	54.3	321	8	ABM80171 Tumour-as
38	44	54.3	347	2	AAV30521 A human m
39	44	54.3	347	4	AAB62698 Human mem
40	44	54.3	347	5	AAB71408 Human HMR
41	44	54.3	347	6	ABU61994 Human pro
42	44	54.3	347	6	ABR58444 Human NOV
43	44	54.3	347	8	ADJ57405 Secretary
44	44	54.3	347	8	ABM80170 Tumour-as
45	44	54.3	347	8	ADP23273 PRO polyp

ALIGNMENTS

RESULT 1
AAG89145
ID AAG89145 standard; protein; 106 AA.
XX
AC AAG89145;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 265.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX (GEST) GENSET.
Dumas Milne Edwards J, Bougueleret L, Jobert S;
WPI; 2001-367870/38.
N-PSDB; AAH64748.

Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Claim 21; Page 806; 921pp; English.

The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patient's own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their

CC expression. The sense and antisense nucleic acids may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and hence to determine which
 CC patients may be in need of restorative therapy. The GENSET polypeptides
 CC may be used as antigens in the production of antibodies and in assays to
 CC identify modulators (agonists and antagonists) of GENSET polypeptide
 CC expression and activity. The present sequence is a GENSET polypeptide of
 CC the invention
 XX
 XX

SQ Sequence 106 AA;

Query Match 100.0%; Score 81; DB 4; Length 106;
 Best Local Similarity 100.0%; Pred. No. 6.2e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
 |||||
 Db 1 MALEVLMLLAVLIWTGA 17

RESULT 2

ABP76237
 ID ABP76237 standard; protein; 143 AA.

XX

AC ABP76237;

XX 21-FEB-2003 (first entry)

XX Human GENSET protein SEQ ID 787.

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;
 KW central nervous system; cardiovascular; gastrointestinal.

XX Homo sapiens.

XX WO200283898-A1.

XX 24-OCT-2002.

XX 18-APR-2001; 2001WO-IB000914.

XX 18-APR-2001; 2001WO-IB000914.

XX (GEST) GENSET.

XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
 DR WPI; 2003-075548/07.

XX New GENSET polynucleotides and polypeptides, useful for treating heavy
 PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
 PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
 PT toxicity.

PS Claim 14; Page 680; 735pp; English.

XX The present invention relates to novel GENSET polynucleotides (AB236404-
 CC AB236911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
 CC and polypeptides are useful in screening and diagnostic assays for
 CC abnormal GENSET expression and/or biological activity. They are also
 CC useful for screening of compounds for treating or preventing GENSET-
 CC related disorders, such as heavy metal toxicity, cancer, inflammatory
 CC diseases, immune disorders, and the neuromuscular, central nervous system
 CC (CNS), cardiovascular or gastrointestinal effects of the toxicity

SQ Sequence 143 AA;

Query Match 100.0%; Score 81; DB 6; Length 143;
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
 |||||
 Db 1 MALEVLMLLAVLIWTGA 17

RESULT 3

ABP76238

ID ABP76238 standard; protein; 148 AA.

XX AC ABP76238;

XX 21-FEB-2003 (first entry)

XX Human GENSET protein SEQ ID 788.

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;
 KW central nervous system; cardiovascular; gastrointestinal.

XX Homo sapiens.

XX WO200283898-A1.

XX 24-OCT-2002.

XX 18-APR-2001; 2001WO-IB000914.

XX 18-APR-2001; 2001WO-IB000914.

XX (GEST) GENSET.

XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
 XX WPI; 2003-075548/07.

XX New GENSET polynucleotides and polypeptides, useful for treating heavy
 PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
 PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
 PT toxicity.

PS Claim 14; Page 680; 735pp; English.

XX The present invention relates to novel GENSET polynucleotides (AB236404-
 CC AB236911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
 CC and polypeptides are useful in screening and diagnostic assays for
 CC abnormal GENSET expression and/or biological activity. They are also
 CC useful for screening of compounds for treating or preventing GENSET-
 CC related disorders, such as heavy metal toxicity, cancer, inflammatory
 CC diseases, immune disorders, and the neuromuscular, central nervous system
 CC (CNS), cardiovascular or gastrointestinal effects of the toxicity

XX Sequence 148 AA;

Query Match 100.0%; Score 81; DB 6; Length 148;
 Best Local Similarity 100.0%; Pred. No. 8.8e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
 |||||
 Db 1 MALEVLMLLAVLIWTGA 17

RESULT 4

AAG89140

ID AAG89140 standard; protein; 158 AA.

XX AC AAG89140;

XX 11-SEP-2001 (first entry)

XX Human secreted protein, SEQ ID NO: 260.

XX

KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
XX Homo sapiens.
OS WO200142451-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-IB001938.
XX
XX 08-DEC-1999; 99US-0169629P.
XX 06-MAR-2000; 2000US-0187470P.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX WPI; 2001-367870/38.
XX N-PSDB; AAH64743.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
XX proteins, useful in gene therapy and vaccination against a variety of
XX diseases, and for diagnosis of those diseases.
XX
XX Claim 21; Page 804; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased GENSET
XX gene expression by rectifying mutations or deletions in a patient's
XX genome that affect the activity of GENSET or by supplementing the
XX nucleic acid molecules of GENSET polypeptides. Conversely, antisense
XX nucleic acid molecules may be administered to down regulate GENSET
XX expression by binding with the cells' own genes and preventing their
XX expression. The sense and antisense nucleic acids may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acid sequences in samples, and hence to determine which
XX patients may be in need of restorative therapy. The GENSET polypeptides
XX may be used as antigens in the production of antibodies and in assays to
XX identify modulators (agonists and antagonists) of GENSET polypeptide
XX expression and activity. The present sequence is a GENSET polypeptide of
XX the invention
XX
XX Sequence 158 AA;
XX
XX Query Match 100.0%; Score 81; DB 4; Length 158;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-05;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MALEVLMLLAVLIWTGA 17
XX |||||
XX 1 MALEVLMLLAVLIWTGA 17
XX
XX
XX RESULT 5
XX ABG97353
XX ID ABG97353 standard; protein; 158 AA.
XX
XX AC ABG97353;
XX
XX 16-DEC-2002 (first entry)
XX
XX Human CGDD4, INCYTE 5284076CD1.
XX
XX Human; cell growth; differentiation; death; CGDD; cancer;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
XX polycythaemia vera; primary thrombocytopaenia; developmental disorder;
XX renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
XX neurological disorder; Alzheimer disease; Parkinson's disease; asthma;

KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
XX Homo sapiens.
XX
XX WO200272830-A2.
XX
XX 19-SEP-2002.
XX
XX 08-FEB-2002; 2002WO-US003715.
XX
XX 09-FEB-2001; 2001US-0268111P.
XX 23-FEB-2001; 2001US-0271175P.
XX 08-MAR-2001; 2001US-0274503P.
XX 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
XX Lu DAM, Richardson TW, Tran UK, Khare R, Walla NK;
XX
XX WPI; 2002-723356/78.
XX N-PSDB; ABS78646.
XX
XX New human proteins associated with cell growth, differentiation and
XX death, useful for diagnosing, treating or preventing autoimmune or
XX inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
XX atherosclerosis or hepatitis.
XX
XX Claim 1; Page 146; 181pp; English.
XX
XX The invention relates to an isolated polypeptide comprising CGDD1-12
XX (cell growth, differentiation and death), a naturally occurring amino
XX acid sequence at least 90% identical to CGDD, a biologically active
XX fragment or an immunogenic fragment. Also included are the
XX polynucleotides encoding CGDD1-12, a recombinant polynucleotide
XX comprising a promoter sequence operably linked to the CGDD
XX polynucleotides, a cell transformed with the recombinant polynucleotide,
XX a transgenic organism comprising the recombinant polynucleotide, an anti-
XX CGDD antibody, screening for compounds which bind to/modulate or are
XX ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
XX CGDD polynucleotide microarray. The polypeptides, polynucleotides,
XX agonists and antagonists are useful for diagnosing, treating or
XX preventing disorders associated with aberrant expression of CGDD,
XX particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
XX cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
XX vera, psoriasis, primary thrombocytopaenia or cancer), developmental
XX disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
XX neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
XX epilepsy), reproductive disorders (e.g. infertility or a disruption in
XX the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
XX (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
XX thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
XX glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
XX osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
XX uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
XX infections. They are also useful in the assessment of the effects of
XX exogenous compounds on the expression of nucleic acid and amino acid
XX sequences of proteins associated with CGDD. The present sequence
XX represents a CGDD protein
XX
XX Sequence 158 AA;
XX
XX Query Match 100.0%; Score 81; DB 5; Length 158;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-05;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MALEVLMLLAVLIWTGA 17

```
Db      1 MALEVLMLLAVLIWTGA 17
|||||
RESULT 6
ABG31324
ID      ABG31324 standard; protein; 158 AA.
XX      AC
XX      ABG31324;
XX      05-NOV-2002 (first entry)
XX      DE
XX      Human GSSP3 polypeptide.
XX      Human; GSSP3; circulating blood glucose level; insulin sensitivity;
XX      body mass; serum glucose regulation; body weight loss; obesity;
XX      metabolic-related disorder; impaired glucose tolerance; stroke;
XX      insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
XX      hypertension; syndrome C; type I diabetes; type II diabetes;
XX      microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
XX      polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
XX      lipotrophy; physical performance; exercise; dyslexia; schizophrenia;
XX      attention-deficit disorder; hyperactivity disorder; psychiatric disorder.
XX      OS
XX      Homo sapiens.
XX      FH
XX      Key      Location/Qualifiers
XX      FT      Peptide      1..17
XX      FT      /label= Signal_peptide
XX      FT      Protein      18..158
XX      FT      /label= Mature_GSSP3_polypeptide
XX      PN      WO200260466-A2.
XX      XX
XX      PD      08-AUG-2002.
XX      XX
XX      PF      01-FEB-2002; 2002WO-IB001333.
XX      XX
XX      PR      02-FEB-2001; 2001US-0266156P.
XX      XX
XX      PA      (GEST ) GENSET.
XX      PI      Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
XX      PI      Bihaun B;
XX      DR      WPI; 2002-608487/65.
XX      DR      N-PSDB; ABK90053.
XX      PT      Reducing circulating glucose levels or increasing insulin sensitivity,
XX      PT      useful for reducing body mass or preventing body weight gain, comprises
XX      PT      administering composition comprising GSSP3 polypeptide.
XX      XX
XX      PS      Claim 1; Page 96-97; 97pp; English.
XX      CC      The present invention relates to the isolation of human GSSP3
XX      CC      polypeptide, and polynucleotide sequences that encode it. The GSSP3
XX      CC      polypeptide reduces circulating blood glucose levels, increases insulin
XX      CC      sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
XX      CC      polynucleotide sequences are useful in serum glucose regulation, fatty
XX      CC      acid metabolism, body weight loss, and prevention of body weight gain.
XX      CC      Compositions comprising GSSP3 polypeptides are useful for controlling
XX      CC      blood glucose levels, for treating metabolic-related diseases or
XX      CC      disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
XX      CC      hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
XX      CC      syndrome C, type I or II diabetes, diabetes related complications,
XX      CC      microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
XX      CC      syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
XX      CC      acanthosis nigricans, leprechaunism and lipotrophy). The polypeptides
XX      CC      are also useful to improve physical performance during work or exercise,
XX      CC      and to treat dyslexia, attention-deficit disorder, attention-
XX      CC      deficit/hyperactivity disorder, and psychiatric disorders such as
XX      CC      schizophrenia. The present sequence represents human GSSP3 polypeptide
XX      XX
```

```
SQ      Sequence 158 AA;
Query Match      100.0%; Score 81; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MALEVLMLLAVLIWTGA 17
Db      1 MALEVLMLLAVLIWTGA 17
|||||
RESULT 7
ABR47931
ID      ABR47931 standard; protein; 158 AA.
XX      AC      ABR47931;
XX      DT      12-JUN-2003 (first entry)
XX      DE      Human secreted protein, SEQ ID 822.
XX      KW      Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
XX      KW      vulnery; antiinflammatory; nootropic; neuroprotective;
XX      KW      antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX      OS      Homo sapiens.
XX      PN      WO200295010-A2.
XX      PD      28-NOV-2002.
XX      PF      19-MAR-2002; 2002WO-US009785.
XX      PR      21-MAR-2001; 2001US-0277340P.
XX      PR      19-JUL-2001; 2001US-0306171P.
XX      PR      13-NOV-2001; 2001US-0331287P.
XX      PA      (HUMA-) HUMAN GENOME SCI INC.
XX      PI      Rosen CA, Ruben SM;
XX      DR      WPI; 2003-129429/12.
XX      PT      Novel human secreted proteins, useful for detecting, preventing,
XX      PT      diagnosing, prognosticating, treating and/or ameliorating cardiovascular
XX      PT      disorders such as arrhythmia.
XX      PS      Claim 13; SEQ ID NO 822; 1881pp; English.
XX      CC      The present invention relates to novel human secreted proteins (ABR47633-
XX      CC      ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
XX      CC      and their coding sequences are useful for the preparation of a diagnostic
XX      CC      or pharmaceutical composition for diagnosing or treating a cardiovascular
XX      CC      disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
XX      CC      arteriosclerosis and myocardial ischaemia), neural disorders, immune
XX      CC      system disorders, muscular disorders, reproductive disorders,
XX      CC      gastrointestinal disorders, pulmonary disorders, renal disorders,
XX      CC      proliferative disorders and/or cancerous diseases and conditions, for
XX      CC      wound healing and epithelial cell proliferation, to treat inflammation or
XX      CC      infection, for treating thrombosis and arteriosclerosis, for treating or
XX      CC      preventing neural damage which occurs in neuronal disorders or
XX      CC      neurodegenerative conditions such as Alzheimer's disease and Parkinson's
XX      CC      disease, to enhance bone and periodontal regeneration and aid in tissue
XX      CC      transplants or bone grafts, to prevent skin aging or hair loss, to
XX      CC      stimulate growth and differentiation of haematopoietic cells and bone
XX      CC      marrow cells when used in combination with other cytokines, to maintain
XX      CC      organs before transplantation or for supporting cell culture of primary
XX      CC      tissues, to increase or decrease differentiation or proliferation of
XX      CC      embryonic stem cells, or to modulate mammalian characteristics or
XX      CC      metabolism. Note: The sequence data for this patent was published in
XX      CC      electronic format and is available from WIPO at
XX      CC      ftp.wipo.int/pub/published_pct_sequences
XX      XX
```

SQ Sequence 158 AA;
Query Match 100.0%; Score 81; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALEVLMLLAVLIWTGA 17
Db 1 MALEVLMLLAVLIWTGA 17
RESULT 8
ID ABR00175
XX ABR00175 standard; protein; 158 AA.
AC ABR00175;
XX
DT 03-APR-2003 (first entry)
XX
DE Human gene 165 encoded secreted protein HTLS08, SEQ ID NO:464.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
XX WO200276488-A1.
XX
XX 03-OCT-2002.
XX
PF 19-MAR-2002; 2002WO-US008276.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
XX
XX N-PSDB; ABZ71354.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
XX preventing, diagnosing, prognosticating, treating and/or ameliorating
XX e.g. gastrointestinal diseases and disorders, or cancers.
XX
XX Claim 13; Page 1048; 1216pp; English.
XX
XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABP00011-ABP00299 represent the proteins they encode.
XX CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
XX CC invention also encompasses antibodies specific for the secreted proteins,
XX CC the use of the secreted proteins in drug screening, and recombinant
XX CC vectors and host cells comprising a nucleic acid of the invention. The
XX CC secreted proteins, nucleic acids encoding them, antibodies or antibody
XX CC fragments specific for the secreted proteins, and modulators of protein
XX CC activity are useful for diagnosing, treating, ameliorating or preventing
XX CC digestive disorders. Such conditions include disorders of the mouth,
XX CC oesophagus, stomach, small intestine, large intestine, liver, biliary
XX CC tract and pancreas, and include cancers of these organs and tissues. The
XX CC secreted proteins and their nucleic acids may also be used in the
XX CC treatment of immune disorders, inflammation, infection,
XX CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX CC of the invention may be used for chromosome identification, chromosome
XX CC mapping, in gene therapy, for identifying individuals from minute
XX CC biological samples, as hybridisation probes, and as molecular weight
XX CC markers. The present sequence represents a human secreted protein of the
XX CC invention

XX
SQ Sequence 158 AA;
Query Match 100.0%; Score 81; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALEVLMLLAVLIWTGA 17
Db 1 MALEVLMLLAVLIWTGA 17
RESULT 9
ID ADB91675
XX ADB91675 standard; protein; 158 AA.
AC ADB91675;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human secreted protein #SEQ ID 621.
XX
KW Secreted protein; gene therapy; antidiabetic; diabetes; human.
XX
OS Homo sapiens.
XX
XX WO2003004622-A2.
XX
XX 16-JAN-2003.
XX
XX 19-MAR-2002; 2002WO-US008124.
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-229407/22.
XX
XX Nucleic acid encoding a human secreted protein is useful in diagnosing or
XX treating diabetes or conditions related to diabetes.
XX
XX Claim 3; SEQ ID NO 621; 1537pp; English.
XX
XX The invention relates to isolated nucleic acid molecules ADB91065-
XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX ADB91834. Also disclosed is a recombinant vector comprising a
XX polynucleotide of the invention, and a recombinant host cell comprising
XX the recombinant vector. The polypeptide of the invention is useful in
XX identifying a binding partner by contacting the polypeptide with a
XX binding partner, and determining whether the binding partner increases or
XX decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX antibody or its fragment, agonist or antagonist are useful for preparing
XX a pharmaceutical composition for diagnosing or treating diabetes or
XX conditions related to diabetes. The present sequence is that of the human
XX immunoglobulin Fc portion used to generate fusion proteins, increasing
XX the stability of the fused protein as compared to the secreted protein
XX only. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 158 AA;
Query Match 100.0%; Score 81; DB 7; Length 158;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALEVLMLLAVLIWTGA 17
Db 1 MALEVLMLLAVLIWTGA 17

RESULT 10	
ADB65520	
ID	ADB65520 standard; protein; 158 AA.
XX	
AC	ADB65520;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	Human protein encoded by clone TEST120282420.
XX	
KW	Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW	cell regeneration; membrane protein; signal transduction-related protein;
KW	transcription-related protein; osteoporosis; neurological disease;
KW	cancer; tumour.
XX	
OS	Homo sapiens.
XX	
PN	EP1308459-A2.
XX	
PD	07-MAY-2003.
XX	
PF	28-MAR-2002; 2002EP-00007401.
XX	
PR	05-NOV-2001; 2001JP-00379298.
XX	
PR	25-JAN-2002; 2002US-00350978.
XX	
PA	(HELI-) HELIX RES INST.
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.
PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI	Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
DR	WPI; 2003-450961/43.
XX	
DR	N-PSDB; ADB63550.
XX	
PT	New polynucleotides and polypeptides, useful for developing a diagnostic
PT	marker or medicines for regulation of their expression and activity, or
PT	as targets of gene therapy.
XX	
PS	Claim 1; Page: 222pp; English.
XX	
CC	The invention discloses a polynucleotide comprising a sequence selected
CC	from 1970 fully defined nucleotide sequences which encode novel
CC	polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC	or its partial peptide, an antibody binding to the polypeptide or peptide
CC	of the polynucleotide, immunologically assaying the polypeptide or
CC	peptide of the polynucleotide by contacting the polypeptide or peptide
CC	with the antibody of the encoded protein, and observing the binding
CC	between the two, a transformant carrying the polynucleotide in an
CC	expressible manner and an antisense polynucleotide. The oligonucleotide
CC	is useful as a primer for synthesising the polynucleotide, or as a probe
CC	for detecting the polynucleotide. The polynucleotides and encoded
CC	proteins are useful as pharmaceutical agents and many disease-related
CC	genes may be included in them, for developing a diagnostic marker or
CC	medicines for regulation of their expression and activity, or as targets
CC	of gene therapy. The genes are involved in tissue and/or cell
CC	regeneration. Membrane proteins, signal transduction-related proteins,
CC	transcription-related proteins, disease-related proteins and genes
CC	encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC	neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC	the activity or expression of the encoded protein to treat diseases. The
CC	sequence presented is a protein of the invention. Note: Some of the
CC	sequence data for this patent is not represented in the printed
CC	specification, but is based on sequence information supplied by the
CC	European Patent Office.
XX	
SQ	Sequence 158 AA;
	Query Match 100.0%; Score 81; DB 7; Length 158;
	Best Local Similarity 100.0%; Pred. No. 9.5e-05;

Query Match 100.0%; Score 81; DB 7; Length 158;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
DB 1 MALEVLMLLAVLIWTGA 17
|||||

RESULT 12
AAB38337
ID AAB38337 standard; protein; 159 AA.
XX AC
AC AAB38337;
DT 31-JAN-2001 (first entry)
XX
XX Human secreted protein encoded by gene 17 clone HTLS08.
DE
XX
XX Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
KW neutropic; antibacterial; virucide; fungicide; ophthalmological; human;
KW vulnary; gene therapy; infection; secreted protein.
XX
XX Homo sapiens.
XX WO200061623-A1.
PN
XX
PD 19-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US008979.
PF
XX 09-APR-1999; 99US-0128693P.
PR
PR 26-APR-1999; 99US-0130991P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;
PI Laflaur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;
PI Young PE;
XX
XX WPI; 2000-647418/62.
DR
XX
XX New nucleic acid molecules encoding 62 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; Page 608-609; 716pp; English.
PS
XX
XX Sequences AAB38321-B38396 represent the amino acid sequences of 62 human
CC secreted proteins encoded by the genes AAC69512-C69587. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of
CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;
CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angiogenesis;
CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections
CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.
CC corneal infection. The polypeptides can also be used to aid wound healing
CC and epithelial cell proliferation, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis
XX
XX Sequence 159 AA;
SQ

Query Match 100.0%; Score 81; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17

Db 1 MALEVLMLLAVLIWTGA 17
|||||

RESULT 13
AD40751
ID AD40751 standard; protein; 80 AA.
XX AC
AC AD40751;
XX
XX 21-APR-2005 (first entry)
DT
XX HIV Vpu protein #160.
DE
XX Immune stimulation; Vpu.
KW
XX Human immunodeficiency virus.
OS
XX WO2005012502-A2.
PN
XX 10-FEB-2005.
PD
XX 29-MAR-2004; 2004WO-US009510.
PF
XX 28-MAR-2003; 2003US-0458026P.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
PI WPI; 2005-132661/14.
DR
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 364-369; 458pp; English.
PS
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HIV Vpu protein used in the scope of the invention.
XX
XX Sequence 80 AA;
SQ

Query Match 56.8%; Score 46; DB 9; Length 80;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWT 15
DB 9 VALVALLIILVIWT 23
:|:|:|:|:|:|:|

RESULT 14
AD40648
ID AD40648 standard; protein; 80 AA.
XX AC
AC AD40648;
XX
XX 21-APR-2005 (first entry)
DT
XX HIV Vpu protein #57.
DE
XX Immune stimulation; Vpu.
KW
XX Human immunodeficiency virus.
XX

PN WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX Disclosure; Page 362-368; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HIV Vpu protein used in the scope of the invention.
XX
XX Sequence 80 AA;
Query Match 55.6%; Score 45; DB 9; Length 80;
Best Local Similarity 40.0%; Pred. No. 20;
Matches 6; Conservative 8; Mismatches 1; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWT 15
Db 10 VALVVVLLIAIVVWT 24
RESULT 15
ADX40735
ID ADX40735 standard; protein; 81 AA.
XX
AC ADX40735;
XX
DT 21-APR-2005 (first entry)
XX
DE HIV Vpu protein #144.
XX
KW Immune stimulation; Vpu.
XX
OS Human immunodeficiency virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
PA
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.

XX Disclosure; Page 364-369; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HIV Vpu protein used in the scope of the invention.
XX
XX Sequence 81 AA;
Query Match 55.6%; Score 45; DB 9; Length 81;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWT 15
Db 10 VALVVVLLIAIVVWT 24
Search completed: May 9, 2006, 10:15:16
Job time : 12.7 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:14:54 ; Search time 4 Seconds
(without alignments)
351.372 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17

Perfect score: 81

Sequence: 1 MALEVLMLLAVLIWTGA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	106	2	US-09-621-976-5223
2	81	100.0	153	2	US-09-621-976-3903
3	81	100.0	158	2	US-09-621-976-5222
4	81	100.0	158	2	US-10-104-047-3674
5	45	55.6	658	2	US-09-252-991A-20495
6	44	54.3	33	2	US-10-290-579A-165
7	44	54.3	86	2	US-10-290-579A-248
8	44	54.3	180	2	US-09-328-352-5855
9	44	54.3	347	1	US-09-004-502-1
10	44	54.3	347	2	US-09-360-125-1
11	43	53.1	33	2	US-10-290-579A-166
12	43	53.1	86	2	US-10-290-579A-245
13	43	53.1	547	2	US-09-489-039A-13843
14	42	51.9	27	2	US-10-290-579A-171
15	42	51.9	28	2	US-10-290-579A-168
16	42	51.9	80	2	US-10-290-579A-240
17	42	51.9	201	2	US-09-302-540-11728
18	42	51.9	446	2	US-09-489-039A-14283
19	41	50.6	28	2	US-10-290-579A-169
20	41	50.6	33	2	US-10-290-579A-157
21	41	50.6	33	2	US-10-290-579A-162
22	41	50.6	316	2	US-09-328-352-8163
23	41	50.6	323	2	US-09-134-001C-4635
24	41	50.6	329	2	US-09-710-279-1942
25	40	49.4	28	2	US-10-290-579A-156
26	40	49.4	33	2	US-10-290-579A-160
27	40	49.4	81	2	US-10-290-579A-247

28	40	49.4	144	2	US-09-270-767-36424	Sequence 36424, A
29	40	49.4	144	2	US-09-270-767-51641	Sequence 51641, A
30	40	49.4	157	2	US-09-489-039A-12693	Sequence 12693, A
31	40	49.4	235	2	US-09-489-039A-8218	Sequence 8218, Ap
32	40	49.4	260	2	US-09-252-991A-23230	Sequence 23230, A
33	40	49.4	277	2	US-09-252-991A-17311	Sequence 17311, A
34	40	49.4	634	2	US-10-104-047-2300	Sequence 2300, Ap
35	39	48.1	242	2	US-09-602-777A-92	Sequence 92, Appl
36	39	48.1	267	2	US-09-602-777A-90	Sequence 90, Appl
37	39	48.1	278	2	US-09-489-039A-9667	Sequence 9667, Ap
38	39	48.1	306	2	US-09-489-039A-10023	Sequence 10023, A
39	39	48.1	537	1	US-08-633-879C-2	Sequence 2, Appli
40	39	48.1	688	2	US-09-489-039A-9813	Sequence 9813, Ap
41	39	48.1	859	2	US-09-538-092-717	Sequence 717, App
42	38.5	47.5	77	2	US-09-749-637A-59	Sequence 59, Appl
43	38	46.9	33	2	US-10-290-579A-158	Sequence 158, App
44	38	46.9	33	2	US-10-290-579A-159	Sequence 159, App
45	38	46.9	38	2	US-09-384-302A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-5223
; Sequence 5223, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5223
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-621-976-5223

Query Match 100.0%; Score 81; DB 2; Length 106;
Best Local Similarity 100.0%; Pred.No. 6.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
|||||
Db 1 MALEVLMLLAVLIWTGA 17

RESULT 2

US-09-621-976-3903
; Sequence 3903, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3903
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens

Qy 1 MALEVLMLAVLIWTGA 17
|||
Db 1 MALEVLMLAVLIWTGA 17

RESULT 7
US-10-290-579A-248
; Sequence 248, Application US/10290579A


```
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 248
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94IN476.104; gene=vpu
US-10-290-579A-248

Query Match          54.3%; Score 44; DB 2; Length 86;
Best Local Similarity 50.0%; Pred. No. 4.9;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 ALEVLMLLAVLIWT 15
Db      16 ALIVALILAIIVMT 29
      |||:::|
      |||:::|

RESULT 8
US-09-328-352-5855
; Sequence 5855, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5855
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5855

Query Match          54.3%; Score 44; DB 2; Length 180;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY      4 EVLMLLAVLIWTG 16
Db      140 KVLISLIIWSG 152
      |||:::|
      |||:::|

RESULT 9
US-09-004-502-1
; Sequence 1, Application US/09004502
; Patent No. 5962263
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA

; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 248
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94IN476.104; gene=vpu
US-10-290-579A-248

Query Match          54.3%; Score 44; DB 1; Length 347;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALEVLMLLAVLIWTG 16
Db      274 AVSVLMLLALLFTG 288
      |:|||||:::|
      |:|||||:::|

RESULT 10
US-09-360-125-1
; Sequence 1, Application US/09360125
; Patent No. 6235715
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/360,125
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,502
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
```

; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0456 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: TONGTUT01
; CLONE: 980615
US-09-360-125-1

Query Match 54.3%; Score 44; DB 2; Length 347;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWTG 16
| : ||||| |::||
DB 274 AVSVLMLLVALLFTG 289

RESULT 11

US-10-290-579A-166
; Sequence 166, Application US/10290579A
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 166
; LENGTH: 33
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: isolate=96ZM651.8; gene=vpv
US-10-290-579A-166

Query Match 53.1%; Score 43; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 2.6;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWT 15
|| | : | : | : |
DB 16 ALIVALLIAIVVWT 29

RESULT 12

US-10-290-579A-245
; Sequence 245, Application US/10290579A
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02

; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 245
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=96ZM651.8; gene=vpv
US-10-290-579A-245

Query Match 53.1%; Score 43; DB 2; Length 86;
Best Local Similarity 50.0%; Pred. No. 7.1;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWT 15
|| | : | : | : |
DB 16 ALIVALLIAIVVWT 29

RESULT 13

US-09-489-039A-13843
; Sequence 13843, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13843
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13843

Query Match 53.1%; Score 43; DB 2; Length 547;
Best Local Similarity 40.0%; Pred. No. 52;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWTG 16
|| | : | : | : |
DB 371 ALAALLVIVLMLWTG 385

RESULT 14

US-10-290-579A-171
; Sequence 171, Application US/10290579A
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 171
; LENGTH: 27
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence of deduced Vpu gene
US-10-290-579A-171

Query Match 51.9%; Score 42; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 3;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWT 15
|||::|:|
Db 10 ALVVTFTIIVVWT 23

RESULT 15
US-10-290-579A-168
; Sequence 168, Application US/10290579A
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: IMMUNODEFICIENCY VIRUS TYPE 1
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 168
; LENGTH: 28
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: Xaa=unknown
; LOCATION: 5
; OTHER INFORMATION: Consensus sequence of deduced Vpu gene
US-10-290-579A-168

Query Match 51.9%; Score 42; DB 2; Length 28;
Best Local Similarity 42.9%; Pred. No. 3.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWT 15
|||::|:|
Db 11 ALVVALLIIVVWT 24

Search completed: May 9, 2006, 10:17:43
Job time : 5.5 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	\$				
1	81	100.0		106	3	US-09-731-872-265	Sequence 265, App
2	81	100.0		106	3	US-09-876-997-265	Sequence 265, App
3	81	100.0		106	5	US-10-643-836-265	Sequence 265, App
4	81	100.0		158	3	US-09-731-872-260	Sequence 260, App
5	81	100.0		158	3	US-09-876-997-260	Sequence 260, App
6	81	100.0		158	4	US-10-050-704-113	Sequence 113, App
7	81	100.0		158	4	US-10-104-047-3674	Sequence 3674, App
8	81	100.0		158	4	US-10-467-535-4	Sequence 4, Appli
9	81	100.0		158	4	US-10-798-512-113	Sequence 113, App
10	81	100.0		158	5	US-10-467-046-3	Sequence 3, Appli
11	81	100.0		158	5	US-10-643-836-260	Sequence 260, App
12	81	100.0		158	5	US-10-472-533-464	Sequence 464, App
13	50	61.7		75	4	US-10-425-115-251583	Sequence 251583,
14	46	56.8		694	4	US-10-156-761-1775	Sequence 11775, A
15	45	55.6		104	3	US-09-833-245-1859	Sequence 1859, Ap
16	45	55.6		104	3	US-09-833-245-1860	Sequence 1860, Ap
17	45	55.6		179	4	US-10-017-161-974	Sequence 974, App
18	45	55.6		570	4	US-10-324-967-34	Sequence 34, Appli
19	45	55.6		963	4	US-10-282-128A-61097	Sequence 61097, A
20	44	54.3		33	6	US-11-135-597-165	Sequence 165, App
21	44	54.3		86	6	US-11-135-597-248	Sequence 248, App
22	44	54.3		306	4	US-10-263-839-160	Sequence 160, App
23	44	54.3		306	4	US-10-262-839-162	Sequence 162, App
24	44	54.3		321	4	US-10-262-839-166	Sequence 166, App
25	44	54.3		347	3	US-09-841-805A-1	Sequence 1, Appli
26	44	54.3		347	3	US-09-790-852-5	Sequence 5, Appli
27	44	54.3		347	4	US-10-263-839-164	Sequence 164, App

; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-876-997-265

Query Match 100.0%; Score 81; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 0.00011; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0;

QY 1 MALEVLMLLAVLIWTGA 17
|||||
Db 1 MALEVLMLLAVLIWTGA 17

RESULT 3

US-10-643-836-265
; Sequence 265, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-643-836-265

Query Match 100.0%; Score 81; DB 5; Length 106;
Best Local Similarity 100.0%; Pred. No. 0.00011; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0;

QY 1 MALEVLMLLAVLIWTGA 17
|||||
Db 1 MALEVLMLLAVLIWTGA 17

RESULT 4

US-09-731-872-260
; Sequence 260, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872

; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-731-872-260

Query Match 100.0%; Score 81; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.00017; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0;

QY 1 MALEVLMLLAVLIWTGA 17
|||||
Db 1 MALEVLMLLAVLIWTGA 17

RESULT 5

US-09-876-997-260
; Sequence 260, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-876-997-260

Query Match 100.0%; Score 81; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.00017; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0;

QY 1 MALEVLMLLAVLIWTGA 17
|||||
Db 1 MALEVLMLLAVLIWTGA 17

RESULT 6

US-10-050-704-113
; Sequence 113, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039F1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18

```
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-113

Query Match      100.0%; Score 81; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALEVLMLLAVLIWTGA 17
Db      1 MALEVLMLLAVLIWTGA 17

RESULT 7
US-10-104-047-3674
; Sequence 3674, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674

Query Match      100.0%; Score 81; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALEVLMLLAVLIWTGA 17
Db      1 MALEVLMLLAVLIWTGA 17

RESULT 8
US-10-467-535-4
; Sequence 4, Application US/10467535
; Publication No. US20040146970A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; YAO, Monique G.
; APPLICANT: ISON, Craig H.; LU, Yan
; APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
; APPLICANT: BAUGHN, Mariah R.; DING, Li
; APPLICANT: XU, Yuming; GIETZEN, Kimberly J.
; APPLICANT: TANG, Tom Y.; LAL, Preeti G.
; APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
; APPLICANT: LU, Dyung Aina M.; RICHARDSON, Thomas W.
; APPLICANT: TRAN, Uyen K.; KHARE, Reena
; APPLICANT: CHAWLA, Narinder K.
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-0903 USN
; CURRENT APPLICATION NUMBER: US/10/467,535
; CURRENT FILING DATE: 2003-08-08
```

```
; PRIOR APPLICATION NUMBER: PCT/US02/03715
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/268,111
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/271,175
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,503
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/274,552
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5284076CD1
US-10-467-535-4

Query Match      100.0%; Score 81; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALEVLMLLAVLIWTGA 17
Db      1 MALEVLMLLAVLIWTGA 17

RESULT 9
US-10-798-512-113
; Sequence 113, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-798-512-113

Query Match      100.0%; Score 81; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALEVLMLLAVLIWTGA 17
Db      1 MALEVLMLLAVLIWTGA 17

RESULT 10
US-10-467-046-3
; Sequence 3, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
```

RESULT 12
US-10-472-533-464
; Sequence 464, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.

Search completed: May 9, 2006, 10:17:00
Job time : 7.4 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:15:47 ; Search time 2.2 Seconds
(without alignments)
357.655 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17
Perfect score: 81
Sequence: 1 MALEVLMLLAVLIWTGA 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /SIDSS/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
5: /SIDSS/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
6: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
7: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
8: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
9: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
10: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
11: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
12: /SIDSS/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	143	9	US-10-475-075-787
2	81	100.0	148	9	US-10-475-075-788
3	81	100.0	158	11	US-11-072-512-3674
4	45	55.6	104	11	US-11-264-096-1859
5	45	55.6	164	11	US-11-264-096-1860
6	45	55.6	163	11	US-11-098-686-10809
7	44	54.3	347	9	US-10-821-234-1379
8	41	50.6	329	9	US-10-793-626-1942
9	41	50.6	467	11	US-11-188-298-3902
10	41	50.6	547	11	US-11-079-463-9820
11	40	49.4	81	9	US-10-853-807A-14
12	40	49.4	313	9	US-10-055-877-234
13	40	49.4	528	11	US-11-188-298-15594
14	40	49.4	634	11	US-11-072-512-2300
15	39	48.1	242	9	US-10-454-437-92
16	39	48.1	267	9	US-10-454-437-90
17	39	48.1	411	11	US-11-188-298-20128
18	39	48.1	527	11	US-11-098-686-10551
19	39	48.1	2767	11	US-11-100-640-38
20	39	48.1	2768	11	US-10-510-101-72
21	38	46.9	138	9	US-10-498-026-43

ALIGNMENTS

RESULT 1
US-10-475-075-787
; Sequence 787, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 787
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-10-475-075-787
Query Match 100.0%; Score 81; DB 9; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWTGA 17
Db 1 MALEVLMLLAVLIWTGA 17
RESULT 2
US-10-475-075-788
; Sequence 788, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

22 38 46.9 292 11 US-11-087-719-73 Sequence 73, Appl
23 38 46.9 292 11 US-11-087-719-75 Sequence 75, Appl
24 38 46.9 303 11 US-11-087-719-71 Sequence 71, Appl
25 38 46.9 304 11 US-11-188-298-7828 Sequence 7828, Ap
26 38 46.9 352 11 US-11-075-047A-34 Sequence 34, Appl
27 38 46.9 364 8 US-10-511-937-2927 Sequence 2927, Ap
28 38 46.9 368 11 US-11-075-047A-22 Sequence 22, Appl
29 38 46.9 370 11 US-11-075-047A-18 Sequence 18, Appl
30 38 46.9 371 11 US-11-075-047A-30 Sequence 30, Appl
31 38 46.9 374 11 US-11-075-047A-24 Sequence 24, Appl
32 38 46.9 380 11 US-11-075-047A-26 Sequence 26, Appl
33 38 46.9 381 9 US-10-793-626-3284 Sequence 3284, Ap
34 38 46.9 389 11 US-11-075-047A-28 Sequence 28, Appl
35 38 46.9 424 9 US-10-453-372-458 Sequence 458, App
36 38 46.9 455 11 US-11-075-047A-32 Sequence 32, Appl
37 38 46.9 471 11 US-11-075-047A-10 Sequence 10, Appl
38 38 46.9 474 11 US-11-075-047A-20 Sequence 20, Appl
39 38 46.9 474 11 US-11-096-568A-12651 Sequence 12651, A
40 38 46.9 477 11 US-11-075-047A-12 Sequence 12, Appl
41 38 46.9 483 11 US-11-075-047A-14 Sequence 14, Appl
42 38 46.9 489 9 US-10-453-372-460 Sequence 460, Appl
43 38 46.9 492 11 US-11-075-047A-16 Sequence 16, Appl
44 38 46.9 515 11 US-11-096-568A-12650 Sequence 12650, A
45 38 46.9 567 11 US-11-096-568A-12649 Sequence 12649, A

```
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 788
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-10-475-075-788

Query Match      100.0%; Score 81; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWTGA 17
      |||||
DB      1 MALEVLMLLAVLIWTGA 17
      |||||

RESULT 3
US-11-072-512-3674
; Sequence 3674, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3674

Query Match      100.0%; Score 81; DB 11; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWTGA 17
      |||||
```

```
Db      1 MALEVLMLLAVLIWTGA 17

RESULT 4
US-11-264-096-1859
; Sequence 1859, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1859
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-264-096-1859

Query Match      55.6%; Score 45; DB 11; Length 104;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWTG 16
      :| | | | | | | |
DB      13 LAFFSLVLISVLLWTG 28

RESULT 5
US-11-264-096-1860
; Sequence 1860, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1860
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-264-096-1860

Query Match      55.6%; Score 45; DB 11; Length 104;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWTG 16
      :| | | | | | | |
DB      13 LAFFSLVLISVLLWTG 28
```

```

RESULT 6
US-11-098-686-10809
; Sequence 10809, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10809
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10809

Query Match      55.6%; Score 45; DB 11; Length 163;
Best Local Similarity 42.9%; Pred. No. 4.7;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY      3 LEVLMMLAVLIWTG 16
Db      144 LQALLIISILYWTG 157
      | : | : : : | | | |
      | : | : : : | | | |

RESULT 7
US-10-821-234-1379
; Sequence 1379, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1379
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1379

Query Match      54.3%; Score 44; DB 9; Length 347;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 ALEVLMMLAVLIWTG 16
Db      274 AVSVLMLLVALLFTG 288
      | : | | | | | : | |
      | : | | | | | : | |

RESULT 8
US-10-793-626-1942
; Sequence 1942, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626

```

```

; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1942
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1942

Query Match      50.6%; Score 41; DB 9; Length 329;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      6 LMLLAVLIWT 15
Db      253 LLIMAILIWT 262
      | : : : | : | | |
      | : : : | : | | |

RESULT 9
US-11-188-298-3902
; Sequence 3902, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 3902
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis ATCC 12228
US-11-188-298-3902

Query Match      50.6%; Score 41; DB 11; Length 467;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      6 LMLLAVLIWT 15
Db      339 LLIMAILIWT 348
      | : : : | : | | |
      | : : : | : | | |

RESULT 10
US-11-079-463-9820
; Sequence 9820, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9820
; LENGTH: 547
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-9820

```

```
Query Match      50.6%; Score 41; DB 11; Length 547;
Best Local Similarity 37.5%; Pred. No. 63;
Matches      6; Conservative      6; Mismatches      4; Indels      0; Gaps      0;

QY      1 MALEVLMLLAVLIWTG 16
Db      463 LSIQIMLYLVLWIG 478

RESULT 11
US-10-853-807A-14
; Sequence 14, Application US/10853807A
; Publication No. US20060034860A1
; GENERAL INFORMATION:
; APPLICANT: Hybrigenics
; TITLE OF INVENTION: Protein-protein interactions in Human Immunodeficiency Virus
; FILE REFERENCE: B5055AA
; CURRENT APPLICATION NUMBER: US/10/853,807A
; CURRENT FILING DATE: 2004-05-26
; PRIOR APPLICATION NUMBER: US 60/333,346
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/385,132
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: PCT/EP 02/13868
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 81
; TYPE: PRT
; ORGANISM: HIV
; FEATURE:
; OTHER INFORMATION: Translation of SEQ ID No7
US-10-853-807A-14

Query Match      49.4%; Score 40; DB 9; Length 81;
Best Local Similarity 40.0%; Pred. No. 14;
Matches      6; Conservative      6; Mismatches      3; Indels      0; Gaps      0;

QY      1 MALEVLMLLAVLIWT 15
Db      10 VALVVATIIAIVVWT 24

RESULT 12
US-10-055-877-234
; Sequence 234, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Rattelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Pattarajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
```

```
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 234
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-055-877-234

Query Match      49.4%; Score 40; DB 9; Length 313;
Best Local Similarity 53.3%; Pred. No. 52;
Matches      8; Conservative      3; Mismatches      4; Indels      0; Gaps      0;

QY      1 MALEVLMLLAVLIWT 15
Db      134 MNWQVCILMAVTIWT 148

RESULT 13
US-11-188-298-15594
; Sequence 15594, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15594
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-11-188-298-15594

Query Match      49.4%; Score 40; DB 11; Length 528;
Best Local Similarity 52.9%; Pred. No. 86;
Matches      9; Conservative      2; Mismatches      6; Indels      0; Gaps      0;

QY      1 MALEVLMLLAVLIWTGA 17
Db      62 LKLEVKTLAKEIWAGS 78
```

RESULT 14

US-11-072-512-2300
; Sequence 2300, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2300
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2300

Query Match 49.4%; Score 40; DB 11; Length 634;
Best Local Similarity 57.1%; Pred. No. 1e+02;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIW 14

|||||:|

Db 95 MRLEALSLLHTLW 108

RESULT 15

US-10-454-437-92
; Sequence 92, Application US/10454437
; Publication No. US20050271115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 92
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-92

Query Match 48.1%; Score 39; DB 9; Length 242;
Best Local Similarity 63.6%; Pred. No. 58;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 VLMLLAVLIWT 15

||:|:|

Db 224 VLLIIAALAWT 234

Search completed: May 9, 2006, 10:17:28

Job time : 2.2 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	50	61.7	334	2	H69148	hypothetical prote
2	47	58.0	274	2	S17804	hypothetical prote
3	45	55.6	101	2	S60434	probable membrane
4	45	55.6	570	2	A83118	probable ATP-bindi
5	45	55.6	653	2	D69815	conserved hypothet
6	44	54.3	347	2	T08826	secretory carrier
7	43	53.1	81	2	T01671	vpu protein - huma
8	43	53.1	81	2	S06409	vpu protein - huma
9	43	53.1	165	2	T28742	hypothetical prote
10	43	53.1	203	2	G85850	hypothetical prote
11	43	53.1	203	2	F64981	hypothetical 22.4
12	43	53.1	203	2	F91006	hypothetical prote
13	43	53.1	752	2	AC3448	nitrogen fixation
14	43	53.1	757	2	C32052	Fixi protein - Rhi
15	43	53.1	757	2	C95344	Fixii copper trans
16	43	53.1	763	2	AG2764	nitrogen fixation
17	43	53.1	763	2	E97545	nitrogen fixation
18	42	51.9	250	2	B82367	conserved hypothet
19	42	51.9	356	2	C70025	multidrug-efflux t
20	42	51.9	443	1	BVECTC	threonine-serine p
21	42	51.9	443	2	C85973	threonine-serine p
22	42	51.9	443	2	D91128	threonine-serine p
23	42	51.9	639	1	F69280	iron (II) transpor
24	42	51.9	1124	2	B84742	probable receptor-
25	41	50.6	81	1	ASLJNK	vpu protein - huma
26	41	50.6	131	2	C72636	hypothetical prote
27	41	50.6	195	2	AG0778	probable membrane
28	41	50.6	249	2	F75583	hypothetical prote
29	41	50.6	256	2	A84946	hypothetical prote

Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALFVLMMLAVLIWTG 16
Db 274 AVSVLMMLVALFTG 288

RESULT 7

T01671
vpu protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 20-Jun-2000
C:Accession: T01671
R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isol
A:Reference number: Z14389; MUID:86245056; PMID:2424612
A:Accession: T01671
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-81 <ALI>
A:Cross-references: UNIPARC:UPI000017865F; EMBL:K03456; NID:G328018
C:Superfamily: HIV-1 vpu protein

Query Match 53.1%; Score 43; DB 2; Length 81;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALEVLMMLAVLIWT 15
Db 10 VALVVTLLIAIVVWT 24

RESULT 8

S06409
vpu protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S06409
R:Cohen, E.A.; Terwilliger, E.F.; Sodroski, J.G.; Haseltine, W.A.
Nature 334, 532-534, 1988
A:Title: Identification of a protein encoded by the vpu gene of HIV-1.
A:Reference number: S06409; MUID:88302445; PMID:3043230
A:Accession: S06409
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-81 <COH>
A:Cross-references: UNIPROT:P05925; UNIPARC:UPI0000138D35
C:Superfamily: HIV-1 vpu protein

Query Match 53.1%; Score 43; DB 2; Length 81;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALFVLMMLAVLIWT 15
Db 11 ALVVAIIAIVVWT 24

RESULT 9

T28742
hypothetical protein F48G7.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28742
R:Clarke, K.; Wohldmann, P.; Harrison, M.
submitted to the EMBL Data Library, January 1998
A:Description: The sequence of C. elegans cosmid F48G7.
A:Reference number: Z20517
A:Accession: T28742
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-165 <CIA>

A:Cross-references: UNIPROT:O44587; UNIPARC:UPI0000079AF2; EMBL:AF039044; PIDN:AAC47948
A:Experimental source: strain Bristol N2; clone F48G7

C:Genetics:
A:Gene: CESP:F48G7.8
A:Map position: 5
A:Introns: 55/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F48G7.8

Query Match 53.1%; Score 43; DB 2; Length 165;
Best Local Similarity 47.1%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MALEVLMMLAVLIWTGA 17
Db 1 MLRLVVVFLTFVILWAGA 17

RESULT 10

G85850
hypothetical protein yohC [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: G85850
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
illier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85850
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <STO>
A:Cross-references: UNIPROT:Q8XEC7; UNIPARC:UPI0000165866; GB:AE005174; NID:G12516443;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yohC

Query Match 53.1%; Score 43; DB 2; Length 203;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MALEVLMMLAVLIW 14
Db 183 LVLEVLLALTIVILW 196

RESULT 11

F64981
hypothetical 22.4 kD protein in pbpg-cdd intergenic region - Escherichia coli (strain K
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: F64981
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F64981
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-203 <BLAT>
A:Cross-references: UNIPROT:Q8XEC7; UNIPARC:UPI00000D0265; GB:AE000303; GB:U000096; NID:
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yohC

Query Match 53.1%; Score 43; DB 2; Length 203;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MALEVLMMLAVLIW 14
Db 183 LVLEVLLALTIVILW 196

```
RESULT 12
F91006
hypothetical protein ECs3022 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F91006
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ohnishi, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F91006
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <HAY>
A:Cross-references: UNIPROT:Q8XBC7; UNIPARC:UPI00001654A6; GB:BA000007; PIDN:BAB36445.1;
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs3022

Query Match 53.1%; Score 43; DB 2; Length 203;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MALEVLMLLAVLIW 14
:||||:|:|:|
Db 183 LVLEVLALTIVILW 196

RESULT 13
AC3448
nitrogen fixation protein fixI (el-e2 type cation ATPase fixi) (EC 3.6.1.-) [imported]
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AC3448
R:DeIvecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AB3252; PMID:11756688
A:Accession: AC3448
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-752 <KUR>
A:Cross-references: UNIPROT:Q8YFF3; UNIPARC:UPI0000058107; GB:AE008917; PIDN:AAL52750.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1569
A:Map position: I
C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d
C:Keywords: hydrolase

Query Match 53.1%; Score 43; DB 2; Length 752;
Best Local Similarity 58.3%; Pred. No. 75;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LMLLAVLIWTGA 17
:||||:|:|:|
Db 128 VMLLSVAVWSGA 139

RESULT 14
C32052
FixI protein - Rhizobium meliloti
C:Species: Rhizobium meliloti
C:Date: 13-Jul-1989 #sequence_revision 13-Jul-1989 #text_change 09-Jul-2004
C:Accession: C32052; S32847; S39994
R:Kahn, D.; David, M.; Domergue, O.; Daveran, M.L.; Ghai, J.; Hirsch, P.R.; Batut, J.
J. Bacteriol. 171, 929-939, 1989
A:Title: Rhizobium meliloti fixGHI sequence predicts involvement of a specific cation pu
A:Reference number: A32052; MUID:89123173; PMID:2536685
```

```
A:Accession: C32052
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <KAH>
A:Cross-references: UNIPROT:P18398; UNIPARC:UPI000012A7F3; EMBL:Z21854; NID:G49403; PIDN:
C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d
C:Keywords: phosphoprotein; transmembrane protein
F:162-495/Domain: ATPase transduction domain homology <ATT>
F:557-699/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 53.1%; Score 43; DB 2; Length 757;
Best Local Similarity 58.3%; Pred. No. 76;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LMLLAVLIWTGA 17
:||||:|:|:|
Db 137 IMLLSVSVWSGA 148

RESULT 15
C95344
FixII copper transport ATPase fixII [imported] - Sinorhizobium meliloti (strain 1021) ma
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95344
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: C95344
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <KUR>
A:Cross-references: UNIPROT:P18398; UNIPARC:UPI000012A7F3; GB:AE006469; PIDN:AAK65317.1;
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Anchors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: fixII
A:Genome: plasmid
C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d

Query Match 53.1%; Score 43; DB 2; Length 757;
Best Local Similarity 58.3%; Pred. No. 76;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LMLLAVLIWTGA 17
:||||:|:|:|
Db 137 IMLLSVSVWSGA 148

Search completed: May 9, 2006, 10:15:41
Job time : 2.9 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:11:54 ; Search time 29.4 Seconds
(without alignments)
407.959 Million cell updates/sec

Title: US-10-664-025-3903_COPY_1_17
Perfect score: 81
Sequence: 1 MALEVLMLLAVLIWTGA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	158	Q8N9U6_HUMAN	Q8N9U6 homo sapien
2	81	100.0	158	Q86WS3_HUMAN	Q86WS3 homo sapien
3	62	76.5	164	Q4FGH8_MOUSE	Q4FGH8 mus musculus
4	53	65.4	174	Q8TGD7_TAESA	Q8TGD7 taenia sagi
5	50	61.7	334	Y378_METTH	O26478 methanobact
6	50	61.7	507	Q4WR18_ASPFU	Q4WR18 aspergillus
7	48	59.3	81	Q8AE43_HIV1	Q8AE43 human immun
8	47	58.0	81	Q8AD88_9HIV1	Q8AD88 human immun
9	47	58.0	219	Q6SNL4_CEBAP	Q6SNL4 cebus apell
10	47	58.0	223	Q6LSD9_PHOPR	Q6LSD9 photobacter
11	47	58.0	274	YPU2_RHOCA	P26158 rhodobacter
12	46	56.8	78	Q8AD96_9HIV1	Q8AD96 human immun
13	46	56.8	80	Q4PU48_9HIV1	Q4PU48 human immun
14	46	56.8	80	Q4PU58_9HIV1	Q4PU58 human immun
15	46	56.8	81	Q5G7F2_9HIV1	Q5G7F2 human immun
16	46	56.8	81	Q8ADY6_9HIV1	Q8ADY6 human immun
17	46	56.8	82	Q7SP8_9HIV1	Q7SP8 human immun
18	46	56.8	230	Q8G5X3_BIFLO	Q8G5X3 bifidobacte
19	46	56.8	694	Q82FL1_STRAW	Q82FL1 streptomyce
20	46	56.8	1152	Q5GYC9_XANOR	Q5GYC9 xanthomonas
21	45	55.6	81	Q41788_9HIV1	Q41788 human immun
22	45	55.6	81	Q58Q01_9HIV1	Q58Q01 human immun
23	45	55.6	81	Q58Q08_9HIV1	Q58Q08 human immun
24	45	55.6	81	Q58Q62_9HIV1	Q58Q62 human immun
25	45	55.6	81	Q5UEH5_9HIV1	Q5UEH5 human immun
26	45	55.6	81	Q6UF05_9HIV1	Q6UF05 human immun
27	45	55.6	81	Q6X6S8_9HIV1	Q6X6S8 human immun
28	45	55.6	81	Q70661_9HIV1	Q70661 human immun
29	45	55.6	81	Q7SV05_9HIV1	Q7SV05 human immun
30	45	55.6	81	Q72L78_9HIV1	Q72L78 human immun
31	45	55.6	81	Q8AD80_9HIV1	Q8AD80 human immun

32	45	55.6	81	2	Q8ADA4_9HIV1	Q8ADA4 human immun
33	45	55.6	81	2	Q8ADC8_9HIV1	Q8ADC8 human immun
34	45	55.6	81	2	Q8ADE4_9HIV1	Q8ADE4 human immun
35	45	55.6	81	2	Q8ADK0_9HIV1	Q8ADK0 human immun
36	45	55.6	81	2	Q8ADL6_9HIV1	Q8ADL6 human immun
37	45	55.6	81	2	Q8ADP8_9HIV1	Q8ADP8 human immun
38	45	55.6	81	2	Q8ADX8_9HIV1	Q8ADX8 human immun
39	45	55.6	81	2	Q8AE51_9HIV1	Q8AE51 human immun
40	45	55.6	81	2	Q8AE75_9HIV1	Q8AE75 human immun
41	45	55.6	81	2	Q8UMDS_9HIV1	Q8UMDS human immun
42	45	55.6	81	2	Q90CW3_9HIV1	Q90CW3 human immun
43	45	55.6	81	2	Q5MGV3_9HIV1	Q5MGV3 human immun
44	45	55.6	81	2	Q4QX90_9HIV1	Q4QX90 human immun
45	45	55.6	82	2	Q6UEX9_9HIV1	Q6UEX9 human immun

ALIGNMENTS

RESULT 1
Q8N9U6_HUMAN
ID Q8N9U6_HUMAN PRELIMINARY; PRT; 158 AA.
AC Q8N9U6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ36198.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX PubMed=14702039; DOI=10.1038/ngl1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirao S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Konai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyaama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuiwa Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK093517; BAC04191.1; -; mRNA.
SQ Ensemble; ENSG00000149507; Homo sapiens.
SQ SEQUENCE 158 AA; 17942 NW; EEB43D70A8A391E8 CRC64;

Query Match 100.0%; Score 81; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. NO. 0.00036;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
 |||||
 DB 1 MALEVLMLLAVLIWTGA 17

RESULT 2

Q86WS3 HUMAN
 ID Q86WS3_HUMAN PRELIMINARY; PRT; 158 AA.

AC Q86WS3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein FLJ36198.
 GN Name=FLJ36198;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Testis;
 RG NIH MGC Project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Testis;
 RG NIH MGC Project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Testis;
 RG NIH MGC Project;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC048121; AAH48121.1; -; mRNA.
 DR EMBL; BC036256; AAH36256.1; -; mRNA.
 DR Ensembl; ENSG00000149507; Homo sapiens.
 KW Hypothetical protein.
 SQ SEQUENCE 158 AA; 17971 MW; EEB43D6FB9AB81FF CRC64;

Query Match 100.0%; Score 81; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 0.00036; Mismatches 0; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
 |||||
 DB 1 MALEVLMLLAVLIWTGA 17

RESULT 3

Q4FZG8 MOUSE
 ID Q4FZG8_MOUSE PRELIMINARY; PRT; 164 AA.

Q4FZG8;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Oocytes;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Oocytes;
 RG NIH MGC Project;
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC099498; AAH99498.1; -; mRNA.
 KW Hypothetical protein.
 SQ SEQUENCE 164 AA; 18849 MW; 49ADE19216BE8606 CRC64;

Query Match 76.5%; Score 62; DB 2; Length 164;
 Best Local Similarity 76.5%; Pred. No. 0.25; Mismatches 2; Indels 0; Gaps 0;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGA 17
 |||||
 DB 1 MALEVLMLLAVLIWTGA 17

RESULT 4

Q8T8D7 TAESA
 ID Q8T8D7_TAESA PRELIMINARY; PRT; 174 AA.

AC Q8T8D7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Taenia saginata (Beef tapeworm).
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Taeniidae; Taenia.
 OX NCBI_TaxID=6206;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Kenyan;
 RA Garate T.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ430566; CAD3242.1; -; mRNA.
 DR InterPro; IPR003961; FN III.
 DR SMART; SM00060; FN3; 1.
 DR PROSITE; PS50853; FN3; 1.

```

KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 174 AA; 18702 MW; 645D04766AC068AF CRC64;

Query Match 55.4%; Score 53; DB 2; Length 174;
Best Local Similarity 75.0%; Pred. No. 5.9;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 LMLLAVLIWTGA 17
   |||||:||||:
Db 4 LMLLALLWTGS 15

RESULT 5
Y378 METHTH STANDARD; PRT; 334 AA.
AC C26478;
DT 30-MAY-2000 (Rel. 39, Created)
DT 10-MAY-2005 (Rel. 39, Last sequence update)
DE Hypothetical UPF0104 protein.
GN OrderedLocusNames=MTH378;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Viacre R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwanji N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT DeltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the UPF0104 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF000823; AAB84884.1; -; Genomic_DNA.
CC PIR; H69148; H69148.1; -; Genomic_DNA.
CC InterPro; IPR005242; CHP374.
CC Pfam; PF03706; UPF0104; 1.
CC TIGRFAMs; TIGR00374; Cons hypoth374; 1.
CC Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 7 27 Potential.
FT TRANSMEM 33 53 Potential.
FT TRANSMEM 120 140 Potential.
FT TRANSMEM 142 162 Potential.
FT TRANSMEM 218 238 Potential.
FT TRANSMEM 247 267 Potential.
FT TRANSMEM 277 297 Potential.
FT TRANSMEM 300 320 Potential.
SQ SEQUENCE 334 AA; 36959 MW; 46D280EC683CB611 CRC64;

Query Match 61.7%; Score 50; DB 1; Length 334;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEVLMLLAVLIWTG 16
   ||:||||:||||
Db 11 ALSILLILLALLIWMG 25

RESULT 6
Q4WR18 ASPFU PRELIMINARY; PRT; 507 AA.
ID Q4WR18;
AC Q4WR18;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Cytochrome P450 monooxygenase, putative.
DE Cytochrome P450 monooxygenase, putative.
GN ORFNames=Afu4914790;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.-Stanley,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman N., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fosker M., Fraser A., Garcia J.L., Garcia M.J., Gobie A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penalta M.A., Perlea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC EMBL; AAHF01000005; BAU9316.1; -; Genomic_DNA.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR002401; EP450I.
CC InterPro; IPR001865; Ribosomal_S2.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR00463; EP450I.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN 1.
CC PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN 1.
CC Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 507 AA; 57810 MW; 16F1D6E6E90612F32 CRC64;

Query Match 61.7%; Score 50; DB 2; Length 507;
Best Local Similarity 53.3%; Pred. No. 38;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWT 15
   |||:|:|:|:|:|
Db 1 MALPILCLAVLIWT 15

RESULT 7
Q8AE43 9HIV1 PRELIMINARY; PRT; 81 AA.
ID Q8AE43;
AC Q8AE43;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Vpu protein.
DE Vpu protein.
GN Name=vpu;

```

```
OS Human immunodeficiency virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus;
OC Primate lentivirus group.
OX NCBI_TaxID=11676;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RX MEDLINE=22375625; PubMed=12487816; DOI=10.1089/089922202320886325;
RA Harris M.E., Serwadda D., Sewankambo N., Kim B., Kigozi G.,
RA Kiwanuka N., Phillips J.B., Wabwire F., Meehen M., Lutalo T.,
RA Lane J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 near full length HIV type 1 genome sequences from Rakai
RT District, Uganda, subtype D and AD recombinants predominate.";
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Harris M.E., Birx D.L., Robb M.L.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Kim B., Phillips J.B., Lane J.R., Merling R., McCutchan F.E.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Lutalo T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Meehen M., Wawer M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Serwadda S., Sewankambo N., Wabwire F., Kigozi G., Kiwanuka N.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF484481; AAN73479.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019076; P:viral release; IEA.
DR InterPro; IPR008187; Vpu.
DR InterPro; IPR009032; Vpu_cyt.
DR Pfam; PF00558; Vpu; 1.
SQ SEQUENCE 81 AA; 9322 MW; 98B7887582935D15 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 81;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALEVIMLLAVLIWT 15
Db 10 VALVALLAIIVWT 24

RESULT 8
Q8AD88_9HIV1
ID Q8AD88_9HIV1 PRELIMINARY; PRT; 81 AA.
AC Q8AD88
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vpu protein.
GN Name=vpu;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus;
OC Primate lentivirus group.
OX NCBI_TaxID=11676;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB38855;
```

```
RX MEDLINE=22375625; PubMed=12487816; DOI=10.1089/089922202320886325;
RA Harris M.E., Serwadda D., Sewankambo N., Kim B., Kigozi G.,
RA Kiwanuka N., Phillips J.B., Wabwire F., Meehen M., Lutalo T.,
RA Lane J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 near full length HIV type 1 genome sequences from Rakai
RT District, Uganda, subtype D and AD recombinants predominate.";
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGK38855;
RA Harris M.E., Birx D.L., Robb M.L.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGK38855;
RA Kim B., Phillips J.B., Lane J.R., Merling R., McCutchan F.E.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGK38855;
RA Meehen M., Wawer M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGK38855;
RA Serwadda S., Sewankambo N., Wabwire F., Kigozi G., Kiwanuka N.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF484520; AAN73822.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019076; P:viral release; IEA.
DR InterPro; IPR008187; Vpu.
DR InterPro; IPR009032; Vpu_cyt.
DR Pfam; PF00558; Vpu; 1.
SQ SEQUENCE 81 AA; 9317 MW; 430257630240FDEC CRC64;

Query Match 58.0%; Score 47; DB 2; Length 81;
Best Local Similarity 53.3%; Pred. No. 26;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALEVIMLLAVLIWT 15
Db 10 VALVALLAIIVWT 24

RESULT 9
Q6SNL4_CEBAP
ID Q6SNL4_CEBAP PRELIMINARY; PRT; 219 AA.
AC Q6SNL4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Olfactory receptor (fragment).
OS Cebus apella (Brown-capped capuchin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini; Cebidae;
OC Cebinae; Cebus.
OX NCBI_TaxID=9515;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gilad Y., Wiebe V., Przeworski M., Lancet D., Paabo S.;
RT "Loss of Olfactory Receptor Genes Coincides with the Acquisition of
RT Full Trichromatic Vision in Primates.";
RL PLOS Biol. 2:0120-0125(2004).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL; AY454839; AARI9450.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
```


DR	GO: 0004384; F:Olfactory receptor activity; IEA.	
DR	GO: 0004872; F:Receptor activity; IEA.	
DR	GO: 0007186; P:G-protein coupled receptor protein signalin. .; IEA.	
DR	GO: 0007165; P:signal transduction; IEA.	
DR	InterPro: IPR000276; GPCR Rhodopn.	
DR	InterPro: IPR000725; Olfact_receptor.	
DR	PANTHER: PTHR11398; Olfact_receptor; 1.	
DR	Pfam: PF00001; 7tm_1; 1.	
DR	PRINTS; PR00245; OLFACTORYR.	
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.	
DR	PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.	
KW	G-protein coupled receptor; Receptor; Transducer; Transmembrane.	
FT	NON_TER	1
FT	NON_TER	219
FT	SEQUENCE	219 AA: 244476 MW: 9B73B54D0446A130 CRC64:
SO		

Query Match	58.0%	Score 47;	DB 2;	Length 219;
Best Local Similarity	56.2%	Pred. No. 55;		
Matches	9;	Conservative	3;	Mismatches 4; Indels 0; Gaps 0;

Qy 1 MALEVLMLLAVLIWTG 16
| : | | : | | : | |
Db 73 MNWKVCMMLAVLWTG 88

```

RESULT 10
Q6LSD9_PROPR Q6LSD9_PROPR PRELIMINARY; PRT; 223 AA.
AC Q6LSD9;
05-JUL-2004 (T-EMBLrel. 27, Created).
05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DT DE Hypothetical protein VV2139.
DE Names=VV2139; OrderedLocusNames=BPBRA1376;
GN Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
NCBI_TaxID=74109;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed:15746425; DOI=10.1126/science.1103341;
RX Vezzi A., Campanaro S., D'Angelo M., Simonati F., Vitulo N.
RA Lauro F.M., Cestaro A., Malacrida G., Simonati B., Cannata
RA Romaldi C., Bartlett D.H., Valle G.;
RT "Life at depth: Photobacterium profundum genome sequence and
RT expression analysis.";
RL Science 307:1459-1461(2005).
RL EMBL; CR378667; CAG19787.1; -; Genomic_DNA.
DR Complete proteome.
KW SEQUENCE 223 AA; 24154 MW; 108FD7606ABFAF87 CRC64;
SQ

```

Query Match 58.0%; Score 47; DB 2; Length 223;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 VLMLLAVLIWTG 16
| | | | : | : | |
Db 181 VLMLVALLVWAG 192

RESULT 11	YFU2 RHQCA	STANDARD;	PRT;	274 AA.
ID	YFU2 RHQCA			
AC	P26158;			
DT	01-MAY-1992	(Rel. 22, Created)		
DT	01-MAY-1992	(Rel. 22, Last sequence update)		
DE	13-SEP-2005	(Rel. 48, Last annotation update)		
DE	Hypothetical 30.4 kDa protein in puHA 5' region (ORF274).			
OS	Rhodobacter capsulatus (Rhodospseudomonas capsulata).			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;			
OC	Rhodobacteraceae; Rhodobacter.			
NCBI_TaxID=1061;				
ORX	[1]			

```

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA Burke D.H., Alberti M., Armstrong G.A., Hearst J.E.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Z11165; CAA77516.1; -; Genomic_DNA.
CC
CC PIR; S17804; S17804.
DR
DR Hypothetical protein; Photosynthesis.
SK
SQ SEQUENCE 274 AA; 30352 MW; 98A650B44B27A01E CRC64;

```

Query Match	58.0%	Score 47;	DB 1;	Length 274;
Best Local Similarity	57.1%	Pred. No. 66;		
Matches	8;	Conservative	3;	Mismatches
				3; Indels 0; Caps 0;

Qy 3 LEVLMLLAVLIWTG 16
Db 123 ISALILCAVLVWTG 136

```

RESULT 12
QBAD96_9HIV1 PRELIMINARY; PRT; 78 AA.
ID QBAD96;
AC QBAD96;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Vpu protein.
GN Name=vpu;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus;
OC Primate lentivirus group.
NCBI_TaxID=11676;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGF03726;
RX MEDLINE=22375625; PubMed=12487816; DOI=10.1089/089822202320886325;
RA Harris M.E., Serwadda D., Sewankambo N., Kim B., Kigozi G.,
RA Kianwuka N., Phillips J.B., Wawere F., Meehen M., Lutalo T.,
RA Lane J.R., Merling R., Gray R., Wawer M., Bix D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 near full length HIV type 1 genome sequences from Rakai
RT District, Uganda, subtype D and AD recombinants predominate.";
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).

```

[2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=99UGF03726;
 RC Harris M.E., Birx D.L., Robb M.L.;
 RA Submitted (FEB-2002) to the ENBL/GenBank/DBJ databases.
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=99UGF03726;
 RC Kim B., Phillips J.B., Lane J.R., Merling R., McCutchan
 RA Submitted (FEB-2002) to the ENBL/GenBank/DBJ databases.

[4]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGF03726;
RA Lutalo T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGF03726;
RA Meehen M., Waver M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGF03726;
RA Serwadda S., Sewankambo N., Wabwire F., Kigozi G., Kiwanuka N.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

<p>DB EMBL; AF484519; AAN73813.1; -; Genomic_DNA. GO; GO:0016020; C:membrane; IEA. GO; GO:0019076; P:viral release; IEA. DR InterPro; IPR008187; Vpu. DR InterPro; IPR009032; Vpu_cyt. DR Pfam; PF00558; Vpu_1; SQ SEQUENCE 78 AA; 8998 MW; 23FA19937BABE989 CRC64;</p>	
<p>Query Match 56.8% ; Score 46; DB 2; Length 78; Best Local Similarity 46.7% ; Pred. No. 35; Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;</p>	
QY 1 MALEVLMMLAVLIWT 15 : Db 10 VALVVALILAIIVWT 24	
<p>RESULT 13 ID Q4PU48_9HIV1 PRELIMINARY; PRT; 80 AA.</p>	
<p>ID Q4PU48; AC AC DT 13-SEP-2005 (TrEMBLrel. 31, Created) DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update) DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)</p>	
<p>DE Vpu protein. GN Name=vpu; OS Human immunodeficiency virus 1. OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae; OC Lentivirus; Primate lentivirus group. OX NCBI_TaxID=11676; RN [1] RP NUCLEOTIDE SEQUENCE.</p>	
<p>RN RC STRAIN=O1YE386; RA Saad M.D., Al-Jauffy A., Graham R.R., Nadai Y., Earhart K.C., RT "HIV from Europe, Africa and Asia Co-circulate in Yemen."; RL Submitted (Oct-2004) to the EMBL/GenBank/DDBJ databases. DR EMBL; AY795903; AAW57833.1; -; Genomic_DNA. DR GO; GO:0019076; P:viral release; IEA. DR InterPro; IPR008187; Vpu. DR Pfam; PF00558; Vpu_1. SQ SEQUENCE 81 AA; 9274 MW; 5CA394AE497C5EC2 CRC64;</p>	
<p>Query Match 56.8% ; Score 46; DB 2; Length 80; Best Local Similarity 46.7% ; Pred. No. 36; Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;</p>	
QY 1 MALEVLMMLAVLIWT 15 : Db 9 VALVVALILAIIVWT 23	
<p>SEARCH completed: May 9, 2006, 10:16:59 Job time : 30.4 secs</p>	

<p>DB EMBL; AF484519; AAN73813.1; -; Genomic_DNA. GO; GO:0016020; C:membrane; IEA. GO; GO:0019076; P:viral release; IEA. DR InterPro; IPR008187; Vpu. DR InterPro; IPR009032; Vpu_cyt. DR Pfam; PF00558; Vpu_1; SQ SEQUENCE 78 AA; 8998 MW; 23FA19937BABE989 CRC64;</p>	
<p>Query Match 56.8% ; Score 46; DB 2; Length 78; Best Local Similarity 46.7% ; Pred. No. 35; Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;</p>	
QY 1 MALEVLMMLAVLIWT 15 : Db 10 VALVVALILAIIVWT 24	
<p>RESULT 13 ID Q4PU48_9HIV1 PRELIMINARY; PRT; 80 AA.</p>	
<p>ID Q4PU48; AC AC DT 13-SEP-2005 (TrEMBLrel. 31, Created) DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update) DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)</p>	
<p>DE Vpu protein. GN Name=vpu; OS Human immunodeficiency virus 1. OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae; OC Lentivirus; Primate lentivirus group. OX NCBI_TaxID=11676; RN [1] RP NUCLEOTIDE SEQUENCE.</p>	
<p>RN RC STRAIN=O4KGH8-1813; RA Cho Y.; RT Submitted (Apr-2005) to the EMBL/GenBank/DDBJ databases. DR EMBL; DQ054370; AAY78574.1; -; Genomic_DNA. DR SQ SEQUENCE 80 AA; 9183 MW; 46FA15083CD0DF09 CRC64;</p>	
<p>Query Match 56.8% ; Score 46; DB 2; Length 80; Best Local Similarity 46.7% ; Pred. No. 36; Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;</p>	
QY 1 MALEVLMMLAVLIWT 15 : Db 9 VALVVALILAIIVWT 23	
<p>SEARCH completed: May 9, 2006, 10:16:59 Job time : 30.4 secs</p>	

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 14:08:32 ; Search time 388.6 Seconds
(without alignments)
7460.154 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103
Perfect score: 51
Sequence: 1 atggcgtagaagtcttgat.....tcttgattggaccggtgct 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sy.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	469	BD109279	BD109279 EST and e
2	51	100.0	469	AR413726	AR413726 Sequence
3	51	100.0	469	AX970560	AX970560 Sequence
4	51	100.0	470	CS072276	CS072276 Sequence
5	51	100.0	477	CQ737424	CQ737424 Sequence
6	51	100.0	497	AX588689	AX588689 Sequence
7	51	100.0	500	AX588688	AX588688 Sequence
8	51	100.0	512	BD107959	BD107959 EST and e
9	51	100.0	512	AR412406	AR412406 Sequence
10	51	100.0	512	AX969240	AX969240 Sequence
11	51	100.0	560	BD110503	BD110503 EST and e
12	51	100.0	560	AR414950	AR414950 Sequence
13	51	100.0	560	AX971784	AX971784 Sequence
14	51	100.0	642	BD109278	BD109278 EST and e
15	51	100.0	642	AR413725	AR413725 Sequence
16	51	100.0	642	AX970559	AX970559 Sequence
17	51	100.0	691	CS072271	CS072271 Sequence
18	51	100.0	1337	AX574405	AX574405 Sequence

19	51	100.0	1480	6	AX600204	AX600204 Sequence
20	51	100.0	1492	6	AX748179	AX748179 Sequence
21	51	100.0	1492	8	AK093517	AK093517 Homo sapi
22	51	100.0	1502	8	BC036256	BC036256 Homo sapi
23	51	100.0	1646	8	BC048121	BC048121 Homo sapi
24	51	100.0	1898	6	BD275948	BD275948 62 Human
25	51	100.0	10115	6	AX574404	AX574404 Sequence
26	51	100.0	100317	14	AP000409	AP000409 Homo sapi
27	51	100.0	101031	14	AP000622	AP000622 Homo sapi
28	51	100.0	176944	8	AP000790	AP000790 Homo sapi
c 29	51	100.0	188367	8	AC013807	AC013807 Homo sapi
30	39	76.5	372	6	BD119756	BD119756 EST and e
31	39	76.5	372	6	AR424203	AR424203 Sequence
32	39	76.5	372	6	AX984897	AX984897 Sequence
33	31.8	62.4	655	9	BC099498	BC099498 Mus muscu
34	31.8	62.4	132494	9	AC126036	AC126036 Mus muscu
35	31.8	62.4	203047	9	AC127289	AC127289 Mus muscu
36	30.8	60.4	245162	14	AC110632	AC110632 Rattus no
c 37	29.4	57.6	171415	14	AP002771	AP002771 Homo sapi
c 38	28	54.9	203635	8	AC148310	AC148310 Pan trogl
39	27.8	54.5	170523	8	AP002387	AP002387 Homo sapi
40	27.6	54.1	165682	14	AC163716	AC163716 Pan trogl
c 41	27.6	54.1	180496	14	AC117328	AC117328 Rattus no
c 42	27.6	54.1	196012	14	AC163727	AC163727 Pan trogl
c 43	27.6	54.1	208880	14	AC128103	AC128103 Rattus no
44	27.6	54.1	208924	14	AC163726	AC163726 Pan trogl
45	27.6	54.1	212868	14	AC163728	AC163728 Pan trogl

ALIGNMENTS

RESULT 1
BD109279
LOCUS BD109279 469 bp linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD109279
VERSION BD109279.1 GI:23204097
KEYWORDS JP 2002010789-A/1356
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 469)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 1356 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/1356
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 10.6999998092651
CC seq VMLLAVLITGCA/EN
FH Key Location/Qualifiers
FT CDS 137..454
FT sig_peptide 137..187.
FEATURES
source
1..469
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```
Query Match      100.0%; Score 51; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 187

RESULT 2
AR413726
LOCUS      AR413726          469 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 1363 from patent US 6639063.
ACCESSION  AR413726
VERSION     AR413726.1  GI:40168836
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 469)
AUTHORS    Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE      EST's and encoded human proteins
JOURNAL    Patent: US 6639063-A 1363 28-OCT-2003;
           Genset S.A.;
           WOX;
FEATURES   Location/Qualifiers
            source
              1..469
              /organism="unknown"
              /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 51; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 187

RESULT 3
AX970560
LOCUS      AX970560          469 bp      DNA      linear      PAT 15-JAN-2004
DEFINITION Sequence 1363 from Patent EP1104808.
ACCESSION  AX970560
VERSION     AX970560.1  GI:40978039
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1
AUTHORS    Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE      ESTs and encoded human proteins
JOURNAL    Patent: EP 1104808-A 1363 06-JUN-2001;
           Genset (FR)
FEATURES   Location/Qualifiers
            source
              1..469
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
              137..>454
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAF15668.1"
              /db_xref="GI:40978040"
              /translation="MALEVLMLLAVLIWTGAENLHVKISCSLDMLMVSVIPVAESRNL
              YIFADELHLMGCPANRIHTYVEFYILVRDGGIRVTRTVIVCKKCMFCQTFMPSI
              KIVF"
CDS
            137..187
            /note="Von Heijne matrix score 10.6999998092651 seq
            VLMLLAVLIWTGA/EN"
sig_peptide
            137..187
            /note="Von Heijne matrix score 10.6999998092651 seq
            VLMLLAVLIWTGA/EN"
```

```
ORIGIN
Query Match      100.0%; Score 51; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 187

RESULT 4
CS072276
LOCUS      CS072276          470 bp      DNA      linear      PAT 05-MAY-2005
DEFINITION Sequence 24 from Patent WO2001042451.
ACCESSION  CS072276
VERSION     CS072276.1  GI:63089511
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1
AUTHORS    Dumas Milne Edwards,J.B., Bougueleret,L. and Jobert,S.
TITLE      Full-length human cDNAs encoding potentially secreted proteins
JOURNAL    Patent: WO 2001042451-A 24 14-JUN-2001;
           Serono Genetics Institute S.A. (FR)
FEATURES   Location/Qualifiers
            source
              1..470
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
              137..457
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAI93414.1"
              /db_xref="GI:63089512"
              /translation="MALEVLMLLAVLIWTGAENLHVKISCSLDMLMVSVIPVAESRNL
              YIFADELHLMGCPANRIHTYVEFYILVRDGGIRVTRTVIVCKKCMFCQTFMPSI
              KIVF"
            137..187
            /note="Von Heijne matrix score 10.7019149919754 seq
            VLMLLAVLIWTGA/EN"
ORIGIN
Query Match      100.0%; Score 51; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 187

RESULT 5
CQ737424
LOCUS      CQ737424          477 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 23358 from Patent WO02068579.
ACCESSION  CQ737424
VERSION     CQ737424.1  GI:42335713
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1
AUTHORS    Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE      Kites, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
            Patent: WO 02068579-A 23358 06-SEP-2002;
            VLMLLAVLIWTGA/EN
JOURNAL
```

```

FEATURES             PE Corporation (NY) (US)
  source              Location/Qualifiers
    1..477
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match          100.0%; Score 51; DB 6; Length 477;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTCTTGTGATTTGGACCGGTGCT 51
    |||||||
Db 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTCTTGTGATTTGGACCGGTGCT 51

RESULT 6
AX588689
LOCUS              497 bp      DNA      linear      PAT 24-JAN-2003
DEFINITION         Sequence 564 from Patent WO02083898.
ACCESSION          AX588689
VERSION            AX588689.1 GI:27900346
KEYWORDS
SOURCE             Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
  1 Bejanin,S., Tanaka,H., Dumas Milne Edwards,J.B., Jobert,S. and
    Giordano,J.Y.
  Full-length human cdnas encoding potentially secreted proteins
  Patent: WO 02083898-A 564 24-OCT-2002;
  GENSET (FR)

FEATURES             Location/Qualifiers
  source
    1..497
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

CDS
    52..>495
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD61628.1"
    /db_xref="GI:27900347"
    /translation="MALEVLMLLAVLIWTGAENLHVKISCLDLMVSVIPVAESRNL
    YIPADELHLGMCPCANRIHTYVYEFILVRDCCGIRTVVSEETLLFQTELYFTPRNID
    HDPEITHLECSSTSRKSVMLTPVSTENEIKLDPSFPFIADPQTAEEL"

sig_peptide
    52..102
    /note="Von Heijne matrix score 10.69999998092651 seq
    VMLLAVLIWTGA/EN"

ORIGIN
Query Match          100.0%; Score 51; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTCTTGTGATTTGGACCGGTGCT 51
    |||||||
Db 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTCTTGTGATTTGGACCGGTGCT 51

RESULT 7
AX588688
LOCUS              500 bp      DNA      linear      PAT 24-JAN-2003
DEFINITION         Sequence 563 from Patent WO02083898.
ACCESSION          AX588688
VERSION            AX588688.1 GI:27900344
KEYWORDS
SOURCE             Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
  1 Bejanin,S., Tanaka,H., Dumas Milne Edwards,J.B., Jobert,S. and
    Giordano,J.Y.
  Full-length human cdnas encoding potentially secreted proteins
  Patent: WO 02083898-A 564 24-OCT-2002;
  GENSET (FR)

FEATURES             Location/Qualifiers
  source
    1..497
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

CDS
    52..>495
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD61628.1"
    /db_xref="GI:27900347"
    /translation="MALEVLMLLAVLIWTGAENLHVKISCLDLMVSVIPVAESRNL
    YIPADELHLGMCPCANRIHTYVYEFILVRDCCGIRTVVSEETLLFQTELYFTPRNID
    HDPEITHLECSSTSRKSVMLTPVSTENEIKLDPSFPFIADPQTAEEL"

sig_peptide
    52..102
    /note="Von Heijne matrix score 10.69999998092651 seq
    VMLLAVLIWTGA/EN"

ORIGIN
Query Match          100.0%; Score 51; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTCTTGTGATTTGGACCGGTGCT 51
    |||||||
Db 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTCTTGTGATTTGGACCGGTGCT 102

RESULT 8
AX588688
LOCUS              512 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION         EST and encoded human protein.
ACCESSION          BD107959
VERSION            BD107959.1 GI:23202777
KEYWORDS           JP 2002010789-A/36.
SOURCE             Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
  1 (bases 1 to 512)
  Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
  EST and encoded human protein
  Patent: JP 2002010789-A 36 15-JAN-2002;
  GENSET CORP

COMMENT
  OS Homo sapiens (human)
  PN JP 2002010789-A/36
  PD 15-JAN-2002
  PF 07-AUG-2000 JP 2002080989
  PR 05-AUG-1999 US 60/147499
  PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
  GIORDANO
  PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
  C12N1/21,
  PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
  C12N15/00
  CC Von Heijne matrix
  CC score 10.6999998092651
  CC seq VMLLAVLIWTGA/EN
  FH Key Location/Qualifiers
  FT CDS 53..511
  FT sig_peptide 53..103.
  FT Location/Qualifiers
    1..512
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"

```

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51
|||||
53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103
|||||

RESULT 9

AR412406 512 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 43 from patent US 6639063.
ACCESSION AR412406
VERSION AR412406.1 GI:40167516
KEYWORDS
SOURCE
ORGANISM

REFERENCE

1 (bases 1 to 512)
Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
EST's and encoded human proteins
TITLE
JOURNAL
Patent: US 6639063-A 43 28-OCT-2003;
Genset S.A.;;
WOX;

FEATURES

source
1..512
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51
|||||
53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103
|||||

RESULT 10

AX969240 512 bp DNA linear PAT 15-JAN-2004
LOCUS
DEFINITION Sequence 43 from Patent EP1104808.
ACCESSION AX969240
VERSION AX969240.1 GI:40975399
KEYWORDS
SOURCE
ORGANISM

REFERENCE

Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

FEATURES

source
1..512
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
53..>511
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF14348.1"
/db_xref="GI:40975400"
/translation="NALVLMILAVLIWTGAENLHKISCSLDMLMVSIVPAESRNL
YIFADELHLMGCPANRHTYVYEFILVRDCGIRTVVSEETLLFQTELYFTPRNID
HDFQEIHLCESTRKSVLTPVSTENEIKLDPSPFIADFTTABELGLLIF"
53..103

CDS

sig_peptide

/note="von Heijne matrix score 10.6999998092651 seq
VLMILAVLIWTGA/EN"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51
|||||
53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103
|||||

RESULT 11

BD110503 560 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION EST and encoded human protein.
ACCESSION BD110503
VERSION BD110503.1 GI:23205321
KEYWORDS JP 2002010789-A/2580.
SOURCE
ORGANISM

REFERENCE

1 (bases 1 to 560)
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
EST and encoded human protein
TITLE
JOURNAL
Patent: JP 2002010789-A 2580 15-JAN-2002;
GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/2580
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FT Key Location/Qualifiers
FT CDS
1..560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source
1..560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51
|||||
53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103
|||||

RESULT 12

AR414950 560 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 2587 from patent US 6639063.
ACCESSION AR414950
VERSION AR414950.1 GI:40170060
KEYWORDS
SOURCE
ORGANISM

REFERENCE

1 (bases 1 to 560)
Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
EST's and encoded human proteins
TITLE

JOURNAL	Patent: US 6639063-A 2587 28-OCT-2003; Genset S.A.;; WOX;	JOURNAL	Patent: JP 2002010789-A 1355 15-JAN-2002; GENSET CORP
FEATURES	source 1..560 Location/Qualifiers /organism="unknown" /mol_type="genomic DNA"	COMMENT	OS Homo sapiens (human) PN JP 2002010789-A/1355 PD 15-JAN-2002 PF 07-AUG-2000 JP 2000280989 PR 05-AUG-1999 US 60/147499 PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI GIORDANO PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC C12N1/21, PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC C12N15/00 CC Von Heijne matrix CC score 10.699998092651 CC seq VMLLAVLIWTGA/EN FH Key Location/Qualifiers FT CDS 53..526 FT sig_peptide 53..103. Location/Qualifiers 1..642
ORIGIN	Query Match 100.0%; Score 51; DB 6; Length 560; Best Local Similarity 100.0%; Pred. No. 2.8e-08; Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FEATURES	source 1..642 Location/Qualifiers /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
	QY 1 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51 Db 53 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103 	ORIGIN	Query Match 100.0%; Score 51; DB 6; Length 642; Best Local Similarity 100.0%; Pred. No. 2.8e-08; Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 13		QY	1 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51 Db 53 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103
LOCUS	AX971784 560 bp DNA linear PAT 15-JAN-2004	RESULT 15	AR413725
DEFINITION	Sequence 2587 from Patent EP1104808.	LOCUS	AR413725 642 bp DNA linear PAT 18-DEC-2003
ACCESSION	AX971784	DEFINITION	Sequence 1362 from patent US 6639063.
VERSION	AX971784.1 GI:40980487	ACCESSION	AR413725
KEYWORDS	Homo sapiens (human)	VERSION	AR413725.1 GI:40168835
SOURCE	Homo sapiens	KEYWORDS	Unknown.
ORGANISM	Homo sapiens	SOURCE	Unknown.
REFERENCE	1 Dumas Milne Edwards, J.B., Jobert, S. and Giordano, J.Y. ESTs and encoded human proteins Patent: EP 1104808-A 2587 06-JUN-2001; Genset (FR)	ORGANISM	Unclassified.
AUTHORS		REFERENCE	1 (bases 1 to 642)
TITLE		AUTHORS	Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
JOURNAL		TITLE	EST's and encoded human proteins
FEATURES	source 1..560 Location/Qualifiers /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 372..>539 /notes="unnamed protein product" /codon_start=1 /protein_id="CAFI6892.1" /db_xref="GI:40980488" /translations="MTLRKSIWECSTSRKSVMLTPVSTENEIKLDPSPIADFQTAA ELGLSSSPNLL"	JOURNAL	Genset S.A.;; WOX;
CDS		FEATURES	source 1..642 Location/Qualifiers /organism="unknown" /mol_type="genomic DNA"
ORIGIN	Query Match 100.0%; Score 51; DB 6; Length 560; Best Local Similarity 100.0%; Pred. No. 2.8e-08; Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	ORIGIN	Query Match 100.0%; Score 51; DB 6; Length 642; Best Local Similarity 100.0%; Pred. No. 2.8e-08; Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	QY 1 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51 Db 53 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103 	QY	1 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 51 Db 53 ATGGCGTTAGAGCTTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCT 103
RESULT 14		Db	
LOCUS	BD109278 642 bp DNA linear PAT 18-SEP-2002	RESULT 14	
DEFINITION	EST and encoded human protein.	LOCUS	BD109278
ACCESSION	BD109278	DEFINITION	EST and encoded human protein.
VERSION	BD109278.1 GI:23204096	ACCESSION	BD109278
KEYWORDS	JP 2002010789-A/1355.	VERSION	BD109278.1 GI:23204096
SOURCE	Homo sapiens (human)	KEYWORDS	JP 2002010789-A/1355.
ORGANISM	Homo sapiens	SOURCE	Homo sapiens (human)
REFERENCE	1 (bases 1 to 642)	ORGANISM	Homo sapiens
AUTHORS	Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
TITLE	EST and encoded human protein	AUTHORS	Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.

This Page Blank (usoft)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:43:18 ; Search time 65.7 Seconds
(without alignments)
5173.506 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 51

Sequence: 1 atggcgctagaagctctgat.....tcttgattggaccggtgct 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	470	5	Aah64748 Human sec
2	51	100.0	497	8	Abz36691 Human GEN
3	51	100.0	500	8	Abz36690 Human GEN
4	51	100.0	691	5	Aah64743 Human sec
5	51	100.0	1337	6	Abk90053 DNA encod
6	51	100.0	1480	6	Abk90053 Human cDN
7	51	100.0	1492	10	Abk90053 Human cDN
8	51	100.0	1898	3	Abk90053 Human sec
9	51	100.0	1898	8	Abk90053 Human sec
10	51	100.0	1898	8	Abk90053 Human sec
11	51	100.0	1898	9	Abk90053 Human sec
12	51	100.0	1898	10	Abk90053 Human sec
13	51	100.0	10115	6	Abk90053 Human GSS
14	25.6	50.2	4802	6	Abk90053 Human pro
15	25	49.0	534	3	Aaf08335 Fusarium
16	25	49.0	534	13	Adu52376 Fusarium
17	25	49.0	534	14	Adz90379 Fusarium
18	24.8	48.6	591	2	Aav69879 Bacillus
19	24.8	48.6	591	2	Aav59478 Bacillus

C	20	24.8	48.6	591	14	ADW24026
	21	24.2	47.5	1179	8	ACD05503
	22	24.2	47.5	1206	10	AD81204
	23	24.2	47.5	1729	10	AD807414
	24	24.2	47.5	2169	12	ADQ84623
	25	24.2	47.5	2215	9	ACC59890
	26	24.2	47.5	4156	10	ADB62163
	27	24.2	47.5	4156	12	ADP10444
	28	24.2	47.5	72149	10	AD81173
C	29	23.4	45.9	4494	6	ABL56459
C	30	23.4	45.9	17758	4	ABL56463
C	31	23.2	45.5	573	4	ABD23488
C	32	23.2	45.5	573	10	ADD66776
C	33	23.2	45.5	573	10	AD88030
	34	23	45.1	682	3	AAF14449
	35	23	45.1	682	13	ADU58490
	36	23	45.1	682	14	AD296493
C	37	22.8	44.7	708	13	ADX34451
C	38	22.8	44.7	1316	12	ADP42140
C	39	22.6	44.3	2067	6	AB214132
	40	22.4	43.9	1079	6	ABN98358
C	41	22.4	43.9	1118	3	AA50472
C	42	22.4	43.9	1120	3	AA50472
C	43	22.4	43.9	1178	5	AA511693
C	44	22.4	43.9	1255	13	ADT14765
	45	22.4	43.9	28000	14	ADZ80847

ALIGNMENTS

RESULT 1

AAH64748

ID AAH64748 standard; cDNA; 470 BP.

XX

AAH64748;

XX

DT 11-SEP-2001 (first entry)

XX

DE Human secreted protein cDNA, SEQ ID NO: 24.

XX

Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

GENSET; ss.

XX

OS Homo sapiens.

XX

PN WO200142451-A2.

XX

PD 14-JUN-2001.

XX

PF 07-DEC-2000; 2000WO-IB001938.

XX

PR 08-DEC-1999; 99US-0169629P.

XX

PR 06-MAR-2000; 2000US-0187470P.

XX

(GEST) GENSET.

XX

Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX

WPI: 2001-367870/38.

XX

P-PSDB; AAG89145.

DR

Full length GENSET human nucleic acids encoding potentially secreted

proteins, useful in gene therapy and vaccination against a variety of

diseases, and for diagnosis of those diseases.

XX

PS Claim 7; Page 583; 921pp; English.

XX

The invention relates to full length GENSET human nucleic acids encoding

potentially secreted proteins. The nucleic acids and the polypeptides

they encode may be used in the prevention, treatment and diagnosis of

diseases associated with inappropriate GENSET gene expression. For

example, they be used to treat disorders associated with decreased GENSET

CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patient's own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention

XX SQ Sequence 470 BP; 137 A; 87 C; 107 G; 139 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 5; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCT 51

Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCT 187

RESULT 2

ABZ36691
ID ABZ36691 standard; cDNA; 497 BP.

XX AC ABZ36691;

XX DT 21-FEB-2003 (first entry)

XX DE Human GENSET coding sequence SEQ ID 564.

XX KW Cytostatic; antiinflammatory; neurotropic; neuroprotective; cardiac;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.

XX OS Homo sapiens.

XX PN WO200283898-A1.

XX PD 24-OCT-2002.

XX PF 18-APR-2001; 2001WO-IB000914.

XX PR 18-APR-2001; 2001WO-IB000914.

XX PA (GEST) GENSET.

XX PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;

XX DR WPI; 2003-075548/07.

XX PT New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.

XX PS Claim 12; Page 579; 735pp; English.

XX CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps

XX

SQ Sequence 497 BP; 127 A; 99 C; 117 G; 154 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 8; Length 497;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCT 51

Db 52 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCT 102

RESULT 3

ABZ36690
ID ABZ36690 standard; cDNA; 500 BP.

XX AC ABZ36690;

XX DT 21-FEB-2003 (first entry)

XX DE Human GENSET coding sequence SEQ ID 563.

XX KW Cytostatic; antiinflammatory; neurotropic; neuroprotective; cardiac;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.

XX OS Homo sapiens.

XX PN WO200283898-A1.

XX PD 24-OCT-2002.

XX PF 18-APR-2001; 2001WO-IB000914.

XX PR 18-APR-2001; 2001WO-IB000914.

XX PA (GEST) GENSET.

XX PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;

XX DR WPI; 2003-075548/07.

XX PT New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.

XX PS Claim 12; Page 578; 735pp; English.

XX CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps

XX SQ Sequence 500 BP; 129 A; 100 C; 113 G; 155 T; 0 U; 3 Other;

Query Match 100.0%; Score 51; DB 8; Length 500;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCT 51

Db 72 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCT 122

RESULT 4

AAH64743
ID AAH64743 standard; cDNA; 691 BP.

XX AAH64743;
AC
XX 11-SEP-2001 (first entry)
DT
XX
DE Human secreted protein cDNA, SEQ ID NO: 19.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
OS Homo sapiens.
XX
XX WO200142451-A2.
PN
XX 14-JUN-2001.
PD
XX
XX 07-DEC-2000; 2000WO-IB001938.
PF
XX 08-DEC-1999; 99US-0169629P.
PR
XX 06-MAR-2000; 2000US-0187470P.
PR
XX (GEST) GENSET.
PA
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX WPI: 2001-367870/38.
XX
DR P-PSDB; AAG89140.
DR
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
XX Claim 7; Page 579; 921pp; English.
PS
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX
SQ Sequence 691 BP; 198 A; 126 C; 151 G; 216 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 5; Length 691;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGGCGTTAGAACTCTTGATGCTCCGCTGCTTCTTGATTGGACCGTGCT 51
|||||
DB 42 ATGGCGTTAGAACTCTTGATGCTCCGCTGCTTCTTGATTGGACCGTGCT 92
|||||
RESULT 5
ABK90053
ID ABK90053 standard; DNA; 1337 BP.
XX
AC ABK90053;
XX
XX 05-NOV-2002 (first entry)
DT
XX DNA encoding human GSSP3 polypeptide.
DE

XX Human; GSSP3; circulating blood glucose level; insulin sensitivity;
KW body mass; serum glucose regulation; body weight loss; obesity;
KW metabolic-related disorder; impaired glucose tolerance; stroke;
KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
KW hypertension; syndrome C; type I diabetes; type II diabetes;
KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
KW lipodystrophy; physical performance; exercise; dyslexia; schizophrenia;
KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
KW gene; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 53..529 /*tag= a
FT CDS /product= "GSSP3 polypeptide"
FT 53..103 /*tag= b
FT sig_peptide
FT 104..526 /*tag= c
FT mat_peptide
FT
XX WO200260466-A2.
PN
XX 08-AUG-2002.
PD
XX 01-FEB-2002; 2002WO-IB001333.
PF
XX 02-FEB-2001; 2001US-0266156P.
PR
XX (GEST) GENSET.
PA
XX Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
PI Bihain B;
PI
XX WPI: 2002-608487/65.
XX P-PSDB; ABG31324.
XX
PT Reducing circulating glucose levels or increasing insulin sensitivity,
PT useful for reducing body mass or preventing body weight gain, comprises
PT administering composition comprising GSSP3 polypeptide.
XX
PS Disclosure; Page 95-96; 97pp; English.
XX
CC The present invention relates to the isolation of human GSSP3
CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
CC polypeptide reduces circulating blood glucose levels, increases insulin
CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and insulin
CC polynucleotide sequences are useful in serum glucose regulation, fatty
CC acid metabolism, body weight loss, and prevention of body weight gain.
CC Compositions comprising GSSP3 polypeptides are useful for controlling
CC blood glucose levels, for treating metabolic-related diseases or
CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
CC syndrome C, type I or II diabetes, diabetes related complications,
CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
CC acanthosis nigricans, leprechaunism and lipodystrophy). The polypeptides
CC are also useful to improve physical performance during work or exercise,
CC and to treat dyslexia, attention-deficit disorder, attention-
CC deficit/hyperactivity disorder, and psychiatric disorders such as
CC schizophrenia. The present sequence encodes human GSSP3 polypeptide
XX
SQ Sequence 1337 BP; 457 A; 214 C; 240 G; 426 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 6; Length 1337;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGGCGTTAGAACTCTTGATGCTCCGCTGCTTCTTGATTGGACCGTGCT 51
|||||
DB 53 ATGGCGTTAGAACTCTTGATGCTCCGCTGCTTCTTGATTGGACCGTGCT 103
|||||

RESULT 6
ABS78646
ID ABS78646 standard; cDNA; 1480 BP.
XX AC
XX ABS78646;
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding, CGDD4, INCYTE 5284076CB1.
XX
XX Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
XX WO200272830-A2.
PN
XX
XX 19-SEP-2002.
PD
XX
XX 08-FEB-2002; 2002WO-US003715.
PF
XX
XX 09-FEB-2001; 2001US-0268111P.
PR
XX 23-FEB-2001; 2001US-0271175P.
PR
XX 08-MAR-2001; 2001US-0274503P.
PR
XX 09-MAR-2001; 2001US-0274552P.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR,
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
XX WPI; 2002-723356/78.
DR
XX P-PSDB; ABG97353.
DR
XX
XX New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
XX Claim 5; Page 169; 181pp; English.
XX
XX The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 1480 BP; 463 A; 244 C; 265 G; 508 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 51; DB 6; Length 1480;
Best Local Similarity 100.0%; Pred. No. 6.4e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATFGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51
Db 29 ATFGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 79
RESULT 7
ADB63550
ID ADB63550 standard; cDNA; 1492 BP.
XX
XX ADB63550;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Human cDNA encoding clone TESTI20282420.
DE
XX
XX Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW tissue regeneration; cell regeneration; membrane protein;
KW signal transduction-related protein; transcription-related protein;
KW osteoporosis; neurological disease; cancer; tumour.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 52.528
FT CDS /*tag= a
FT /product= "Clone TESTI20282420 protein"
XX
XX EP1308459-A2.
PN
XX
XX 07-MAY-2003.
PD
XX
XX 28-MAR-2002; 2002EP-00007401.
PF
XX
XX 05-NOV-2001; 2001JP-00379298.
PR
XX 25-JAN-2002; 2002US-00350978.
PR
XX
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
XX WPI; 2003-450961/43.
DR
XX P-PSDB; ADB65520.
DR
XX New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
XX Claim 1; Page; 222pp; English.
XX
XX The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide

CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX
 SQ Sequence 1492 BP; 452 A; 250 C; 272 G; 518 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 10; Length 1492;
 Best Local Similarity 100.0%; Pred. No. 6.5e-10;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGCTTTGATGCTCCCTCGCTGCTTGTGATTGGACCGTGCT 51
 |||||
 Db 52 ATGGCGTTAGAGCTTTGATGCTCCCTCGCTGCTTGTGATTGGACCGTGCT 102
 |||||

RESULT 8
 AAC69528
 ID AAC69528 standard; DNA; 1898 BP.

XX AAC69528;

XX 31-JAN-2001 (first entry)

DE Human secreted protein gene 17 clone HTEL508.

XX Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
 KW cytotatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
 KW nootropic; antibacterial; virucide; fungicide; ophthalmological;
 KW vulnerary; gene therapy; infection; human; secreted protein; ss.

XX Homo sapiens.

XX WO200061623-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US008979.

XX 09-APR-1999; 99US-0128693P.

XX 26-APR-1999; 99US-0130991P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;

PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;

XX Young PB;

DR WPI; 2000-647418/62.

XX New nucleic acid molecules encoding 62 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.

XX Claim 1; Page 543-544; 716pp; English.

XX The invention relates to the isolation of genes (AAC69512-C69587)

CC encoding 62 human secreted proteins (AAB38321-B38396). The genes can be

CC used to generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion (AAC69503) for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
 CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of
 CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;
 CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angiogenesis;
 CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections
 CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.
 CC corneal infection. The polypeptides can also be used to aid wound healing
 CC and epithelial cell proliferation, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis
 XX
 SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Query Match 100.0%; Score 51; DB 3; Length 1898;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGCTTTGATGCTCCCTCGCTGCTTGTGATTGGACCGTGCT 51
 |||||
 Db 15 ATGGCGTTAGAGCTTTGATGCTCCCTCGCTGCTTGTGATTGGACCGTGCT 65
 |||||

RESULT 9

AAC50642

ID AAC50642 standard; cDNA; 1898 BP.

XX AAC50642;

XX 12-JUN-2003 (first entry)

DE Human secreted protein coding sequence, SEQ ID 309.

XX Cardiant; antiarthritic; antiarteriosclerotic; vasotropic; cytostatic;
 KW vulnerary; antiinflammatory; nootropic; neuroprotective;
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder; gene; ss.

XX Homo sapiens.

XX WO200295010-A2.

XX 28-NOV-2002.

XX 19-MAR-2002; 2002WO-US009785.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-129429/12.

PT Novel human secreted proteins, useful for detecting, preventing,
 PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
 PT disorders such as arrhythmia.

XX Claim 21; SEQ ID NO 309; 1881pp; English.

XX The present invention relates to novel human secreted proteins (ABR47633-
 CC ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
 CC and their coding sequences are useful for the preparation of a diagnostic
 CC or pharmaceutical composition for diagnosing or treating a cardiovascular
 CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
 CC arteriosclerosis and myocardial ischemia), neural disorders, immune

CC system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism. Note: The sequence data for this patent was published in
 CC electronic format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Query Match 100.0%; Score 51; DB 8; Length 1898;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGTCTGCTTGGATTGGACCGGTGCT 51
 |||||
 Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGTCTGCTTGGATTGGACCGGTGCT 65

RESULT 10
 ABZ71354
 ID ABZ71354 standard; cDNA; 1898 BP.
 AC ABZ71354;
 XX
 XX 04-APR-2003 (first entry)
 XX
 XX Secreted protein-encoding gene 165 cDNA clone HTL808, SEQ ID NO:175.
 XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic;
 KW antiinflammatory; immunosuppressive; vulnery; gene therapy; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200276488-A1.
 XX
 XX 03-OCT-2002.
 XX
 XX 19-MAR-2002; 2002WO-US008276.
 XX
 XX 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2003-029900/02.
 DR P-PSDB; ABR00175.
 DR
 XX New human secreted proteins and nucleic acids, useful for detecting,
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 PT e.g. gastrointestinal diseases and disorders, or cancers.
 XX
 XX Claim 21; Page 853-854; 1216pp; English.
 PS
 XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC

CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute
 CC biological samples, as hybridisation probes, and as molecular weight
 CC markers. The present sequence represents a human secreted protein-
 CC encoding cDNA clone of the invention
 XX
 SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Query Match 100.0%; Score 51; DB 8; Length 1898;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGTCTGCTTGGATTGGACCGGTGCT 51
 |||||
 Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGTCTGCTTGGATTGGACCGGTGCT 65

RESULT 11
 ADB91291
 ID ADB91291 standard; cDNA; 1898 BP.
 AC ADB91291;
 XX
 XX 04-DEC-2003 (first entry)
 XX
 XX Human secreted protein cDNA #SEQ ID 237.
 DE
 XX Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.
 KW
 XX Homo sapiens.
 OS
 XX WO2003004622-A2.
 PN
 XX 16-JAN-2003.
 PD
 XX 19-MAR-2002; 2002WO-US008124.
 PF
 XX 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2003-229407/22.
 DR
 XX Nucleic acid encoding a human secreted protein is useful in diagnosing or
 PT treating diabetes or conditions related to diabetes.
 PT
 XX Claim 9; SEQ ID NO 237; 1537pp; English.
 PS
 XX The invention relates to isolated nucleic acid molecules ADB91065-
 CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
 CC ADB91834. Also disclosed is a recombinant vector comprising a
 CC polynucleotide of the invention, and a recombinant host cell comprising
 CC the recombinant vector. The polypeptide of the invention is useful in
 CC identifying a binding partner by contacting the polypeptide with a
 CC binding partner, and determining whether the binding partner increases or
 CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
 CC

CC antibody or its fragment, agonist or antagonist are useful for preparing
 CC a pharmaceutical composition for diagnosing or treating diabetes or
 CC conditions related to diabetes. The present sequence is that of the human
 CC immunoglobulin Fc portion used to generate fusion proteins, increasing
 CC the stability of the fused protein as compared to the secreted protein
 CC only. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Query Match 100.0%; Score 51; DB 9; Length 1898;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGGTTAGAGCTTTGATGCTCTCGCTGCTCTTGTGATTGGACCGGTGCT 51

Db 15 ATGCGGTTAGAGCTTTGATGCTCTCGCTGCTCTTGTGATTGGACCGGTGCT 65

RESULT 12

ADCT73723

ID ADCT73723 standard; DNA; 1898 BP.

XX AC ADCT73723;

XX DT 01-JAN-2004 (first entry)

XX DE Human secreted protein-related DNA - SEQ ID 356.

XX KW antianaemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
 KW antidiabetic; immunosuppressive; dermatologic; nephrotropic;
 KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
 KW fungicide; antiparasitic; antiarteriosclerotic; vulnary; cytostatic;
 KW haemopoietic; haematologic; anaemia; autoimmune disorder;
 KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
 KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
 KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
 KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
 KW human; gene; ds.

XX OS Homo sapiens.

XX PN WO2003038063-A2.

XX PD 08-MAY-2003.

XX PF 19-MAR-2002; 2002WO-US008277.

XX PR 21-MAR-2001; 2001US-0277340P.

XX PR 19-JUL-2001; 2001US-0306171P.

XX PR 13-NOV-2001; 2001US-0331287P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Rosen CA, Ruben SM;

XX XX WPI; 2003-430516/40.

XX XX P-PSDB; ADCT74338.

XX XX New human secreted polypeptide for diagnosing, preventing or treating
 PT hemopoietic or hematologic disorders (e.g. anemia), autoimmune
 PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
 PT atherosclerosis).

XX PS Claim 27; SEQ ID NO 356; 2272pp; English.

XX XX The invention relates to a novel human secreted polypeptide comprising a
 CC defined sequence given in the specification. The polypeptide, nucleic
 CC acid molecule, antibody, agonist or antagonist of the invention may be
 CC useful for preparing a composition for diagnosing or treating a
 CC haemopoietic or haematologic disorder such as anaemia, autoimmune
 CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,

CC diabetes, systemic lupus erythematosus or glomerulonephritis,
 CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
 CC disease, wounds and hyperproliferative disorders including
 CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
 CC parasitic infections. The polypeptide may also be used during gene
 CC therapy procedures and for identifying a binding partner by contacting
 CC the polypeptide with a binding partner and determining whether the
 CC binding partner increases or decreases the activity of the polypeptide.
 CC The current sequence is that of the human secreted protein-related DNA of
 CC the invention.

XX SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Query Match 100.0%; Score 51; DB 10; Length 1898;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGGTTAGAGCTTTGATGCTCTCGCTGCTCTTGTGATTGGACCGGTGCT 51

Db 15 ATGCGGTTAGAGCTTTGATGCTCTCGCTGCTCTTGTGATTGGACCGGTGCT 65

RESULT 13

ABK90052

ID ABK90052 standard; DNA; 10115 BP.

XX AC ABK90052;

XX DT 05-NOV-2002 (first entry)

XX XX Human GSSP3 genomic sequence.

XX KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
 KW body mass; serum glucose regulation; body weight loss; obesity;
 KW metabolic-related disorder; impaired glucose tolerance; stroke;
 KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
 KW hypertension; syndrome C; type I diabetes; type II diabetes;
 KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
 KW polycystic ovarian syndrome; acanthosis nigrican; leprechaunism;
 KW lipotrophy; physical performance; exercise; dyslexia; schizophrenia;
 KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
 KW gene; ds.

XX OS Homo sapiens.

XX XX WO200260466-A2.

XX XX 08-AUG-2002.

XX XX 01-FEB-2002; 2002WO-IB001333.

XX XX 02-FEB-2001; 2001US-0266156P.

XX XX (GEST) GENSET.

XX XX Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
 PI Bihain B;

XX XX WPI; 2002-608487/65.

XX XX Reducing circulating glucose levels or increasing insulin sensitivity,
 PT useful for reducing body mass or preventing body weight gain, comprises
 PT administering composition comprising GSSP3 polypeptide.

XX PS Disclosure; Page 90-95; 97pp; English.

XX XX The present invention relates to the isolation of human GSSP3
 CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
 CC polypeptide reduces circulating blood glucose levels, increases insulin
 CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
 CC polynucleotide sequences are useful in serum glucose regulation, fatty
 CC acid metabolism, body weight loss, and prevention of body weight gain.
 CC Compositions comprising GSSP3 polypeptides are useful for controlling

CC blood glucose levels, for treating metabolic-related diseases or
 CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
 CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
 CC syndrome C, type I or II diabetes, diabetes related complications,
 CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
 CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
 CC acanthosis nigricans, leprechaunism and lipodystrophy). The polypeptides
 CC are also useful to improve physical performance during work or exercise,
 CC and to treat dyslexia, attention-deficit disorder, attention-
 CC deficit/hyperactivity disorder, and psychiatric disorders such as
 CC schizophrenia. The present sequence encodes human GSP3 polypeptide
 XX Sequence 10115 BP; 2860 A; 1941 C; 1815 G; 3475 T; 0 U; 24 Other;
 SQ Query Match 100.0%; Score 51; DB 6; Length 10115;
 Best Local Similarity 100.0%; Pred. No. 1e-09; Mismatches 0; Gaps 0;
 Matches 51; Conservative 0; Indels 0; Indels 0; Gaps 0;
 QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCT 51
 DB 2439 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCT 2489
 RESULT 14
 AAD46955
 ID AAD46955 standard; DNA; 4802 BP.
 AC AAD46955;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Human protein-encoding gene 4, SEQ IN NO:25.
 KW Human; immune system disorder; HIV infection; myocardial infarction;
 KW human immunodeficiency virus; arthritis; hyperproliferative disease;
 KW acquired immune deficiency syndrome; inflammation; asthma; myopathy;
 KW allergy; cancer; cardiac oedema; duodenal ulcer; psoriasis; sepsis;
 KW neuromuscular system disorder; multiple myeloma; pulmonary disorder;
 KW cardiovascular disorder; rhabdomyosarcoma; gastrointestinal disorder;
 KW multiple sclerosis; immune-mediated thrombocytopaenia; myocarditis;
 KW leiomyosarcoma; autoimmune disorder; Crohn's disease; gene therapy;
 KW hyperaemia; AIDS; infection; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT exon 1..527
 FT intron 528..1441
 FT exon 1442..1776
 FT intron 1777..2804
 FT exon 2805..2986
 FT intron 2987..3542
 FT exon 3543..3665
 FT intron 3666..3915
 FT exon 3916..4802
 FT intron 4803..5115
 PN WO200272763-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US006990.
 XX
 PR 09-MAR-2001; 2001US-0274214P.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX Shi Y, Ni J, Ruben SM;
 XX WPI; 2002-759842/82.
 XX
 XX New proteins encoded by genes contained in cDNA clones (e.g. HAPQL38 or
 PT HHFOS77), useful for preventing, treating, ameliorating or diagnosing
 PT e.g. AIDS, sepsis, brain cancer, Crohn's disease or myocardial
 PT infarction.
 XX
 PS Disclosure; Page 422-423; 429pp; English.
 XX
 CC AAD46949-AAD46954 represent cDNAs corresponding to novel human protein
 CC genes, and AAE29290-AAE29295 represent the proteins they encode. AAD46955
 CC AAD46963 represent novel human protein genes and AAE29296-AAE29297
 CC represent novel human protein fragments. Sequences of the invention are
 CC useful for preventing, treating or ameliorating medical conditions or for
 CC wound healing. These conditions include diseases or disorders of the
 CC immune system (e.g. HIV infection, autoimmune disorders, arthritis,
 CC asthma, AIDS, sepsis, psoriasis or inflammation) or neuromuscular system
 CC (e.g. multiple sclerosis, myocarditis, or myopathies), disorders
 CC associated with Fc receptor binding by antibody (e.g. immune-mediated
 CC thrombocytopaenia, inflammatory responses or allergic responses),
 CC hyperproliferative diseases (e.g. multiple myeloma, rhabdomyosarcoma,
 CC lung cancer, brain cancer or leiomyosarcoma), gastrointestinal disorders
 CC (e.g. Crohn's disease or duodenal ulcers), pulmonary disorders,
 CC infections or cardiovascular disorders (e.g. hyperaemia, myocardial
 CC infarction or cardiac oedema). The invention is useful in gene therapy.
 CC The present sequence represents human protein gene of the invention
 XX
 SQ Sequence 4802 BP; 1073 A; 1390 C; 1326 G; 1013 T; 0 U; 0 Other;
 Query Match 50.2%; Score 25.6; DB 6; Length 4802;
 Best Local Similarity 70.8%; Pred. No. 10;
 Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 2 TGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTG 49
 DB 3214 TGGCTTGGGATCTTTAAGTCTCTCAGTTTGACATGGCCAGGTG 3261
 RESULT 15
 AAF08335
 ID AAF08335 standard; cDNA; 534 BP.
 XX
 AC AAF08335;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Fusarium venenatum EST SEQ ID NO:858.
 XX
 KW Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 XX
 OS Fusarium venenatum.
 XX
 PN WO200056762-A2.
 XX
 PD 28-SEP-2000.
 XX
 PF 22-MAR-2000; 2000WO-US007781.
 XX
 PR 22-MAR-1999; 99US-00273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 XX

WPI; 2000-594572/56.

Monitoring differential expression of genes in filamentous fungal cells uses fluorescence-labeled nucleic acids isolated from the cells and a substrate of expressed sequence tags.

Claim 86; Page 711; 3161pp; English.

The present invention describes a method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene products to facilitate analysis of the results. AAF07478 to AAF11247 represents ESTs from *Fusarium venenatum*; AAF1248 to AAF11853 represents ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

Sequence 534 BP; 132 A; 148 C; 103 G; 130 T; 0 U; 21 Other;

```

every Match      49.0%; Score 25; DB 3; Length 534;
1st Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGTCTTGATTTGG 42

79 ATGTCGNTCAATGTCCTTGATAATGCTCGCTTCTTCACATTGG 120

Search completed: May 9, 2006, 14:08:20
Job time : 68.7 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:55:51 ; Search time 16.3 Seconds
(without alignments)
5561.696 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 51

Sequence: 1 atggcgtagaagcttgat.....tcttgattggaccggtgct 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
 - 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
 - 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
 - 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	459	3	US-09-621-976-1363
2	51	100.0	512	3	US-09-621-976-43
3	51	100.0	560	3	US-09-621-976-2587
4	51	100.0	642	3	US-09-621-976-1362
5	51	100.0	1492	3	US-10-104-047-1704
6	39	76.5	372	3	US-09-621-976-15700
7	25	49.0	534	3	US-09-533-559-858
8	24.8	48.6	591	3	US-09-402-668-1
9	24.2	47.5	4156	3	US-10-104-047-317
10	23.2	45.5	573	3	US-09-854-133-468
11	23	45.1	682	3	US-09-533-559-6972
12	23	45.1	19412	3	US-09-949-016-17094
13	22.2	43.5	1722	3	US-09-434-840-1
14	22.2	43.5	1722	3	US-09-434-840-3
15	22.2	43.5	1731	3	US-09-434-840-62
16	22.2	43.5	1731	3	US-09-434-840-54
17	22.2	43.5	1732	3	US-09-434-840-56
18	22.2	43.5	1732	3	US-09-434-840-60
19	22.2	43.5	1733	3	US-09-434-840-58
20	22.2	43.5	11168	3	US-09-434-840-5
21	22.2	43.5	13023	3	US-09-949-016-16292
22	22	43.1	152393	3	US-09-949-016-14514
23	22	43.1	152393	3	US-09-949-016-14515
24	22	43.1	156894	3	US-09-949-016-12765

ALIGNMENTS

RESULT 1

US-09-621-976-1363
; Sequence 1363, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1363
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-621-976-1363

Query Match 100.0%; Score 51; DB 3; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCTCGCTCTTGTGATTTGGACCGGTGCT 51

Db 137 ATGGCGTTAGAGTCTTGATGCTCTCTCGCTCTTGTGATTTGGACCGGTGCT 187

RESULT 2

US-09-621-976-43
; Sequence 43, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2

Sequence 12766, A
Sequence 16957, A
Sequence 16958, A
Sequence 16959, A
Sequence 1, Appli
Sequence 2813, Ap
Sequence 2, Appli
Sequence 1, Appli
Sequence 10565, A
Sequence 1, Appli
Sequence 2, Appli
Sequence 16072, A
Sequence 16072, A
Sequence 2524, Ap
Sequence 17806, A
Sequence 1532, Ap
Sequence 21817, A
Sequence 21818, A
Sequence 21819, A
Sequence 21820, A
Sequence 83735, A

; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 43
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..511
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-621-976-43

Query Match 100.0%; Score 51; DB 3; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 103

RESULT 3

US-09-621-976-2587
; Sequence 2587, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2587
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 372..539
US-09-621-976-2587

Query Match 100.0%; Score 51; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 103

RESULT 4

US-09-621-976-1362
; Sequence 1362, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm

; SEQ ID NO 1362
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..526
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-621-976-1362

Query Match 100.0%; Score 51; DB 3; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 103

RESULT 5

US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1704

Query Match 100.0%; Score 51; DB 3; Length 1492;
Best Local Similarity 100.0%; Pred. No. 3.5e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 51
|||||
DB 52 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCT 102

RESULT 6

US-09-621-976-15700
; Sequence 15700, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15700
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 351


```
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(573)
; OTHER INFORMATION: n = A,T,C or G
US-09-854-133-468

Query Match          45.5%; Score 23.2; DB 3; Length 573;
Best Local Similarity 75.7%; Pred. No. 19;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 15 CTTGATGCTCCTCGCTGCTTGTGATTTGGACCGGTGCT 51
      ||||| ||||| ||||| ||||| ||||| |||||
Db 29 CTTGATGCTTGTCACTGTGATGATGTGGCCCTNGCT 65

RESULT 11
US-09-533-559-6972
; Sequence 6972, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6972
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(682)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-6972

Query Match          45.1%; Score 23; DB 3; Length 682;
Best Local Similarity 68.1%; Pred. No. 23;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2 TGGCGTTAGAGTCTTTGATGCTCCTCGCTGCTTGTGATTTGGACCGGT 48
      ||||| ||||| ||||| ||||| ||||| |||||
Db 600 TGGCGCTACATCCTATGATGCTATTTCATGCTCTTGTGATCTGGACCGT 646

RESULT 12
US-09-949-016-17094
; Sequence 17094, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 17094
; LENGTH: 19412
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17094

Query Match          45.1%; Score 23; DB 3; Length 19412;
Best Local Similarity 68.1%; Pred. No. 49;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTTGATGCTCCTCGCTGCTTGTGATTTGGACCGG 47
      ||||| ||||| ||||| ||||| ||||| |||||
Db 13287 ATGCTTTAAATCCATGCTGCTCCTCGCTACCTGGAGCTGCAGCAG 13333

RESULT 13
US-09-434-840-1/c
; Sequence 1, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootlie, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)..(1616)
US-09-434-840-1

Query Match          43.5%; Score 22.2; DB 3; Length 1722;
Best Local Similarity 64.7%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTTGATGCTCCTCGCTGCTTGTGATTTGGACCGGTCT 51
      ||||| ||||| ||||| ||||| ||||| |||||
Db 695 ATAGTTCGGTACCTCTGATGTTCTCGGTGTTTGGTGTGCTGTGTTGTT 645

RESULT 14
US-09-434-840-3/c
; Sequence 3, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootlie, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)..(1064)
US-09-434-840-3

Query Match      43.5%; Score 22.2; DB 3; Length 1722;
Best Local Similarity 64.7%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTTGGACCGGTGCT 51
      |||||
Db      695 ATAGTGCCGTACCTCTGATGTTCCTCGGTGTGTTTGAGTTGCTGTGTT 645

RESULT 15
US-09-434-840-62/c
; Sequence 62, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootle, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1548)
US-09-434-840-62

Query Match      43.5%; Score 22.2; DB 3; Length 1731;
Best Local Similarity 64.7%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTTGGACCGGTGCT 51
      |||||
Db      705 ATAGTGCCGTACCTCTGATGTTCCTCGGTGTGTTTGAGTTGCTGTGTT 655
```

Search completed: May 9, 2006, 13:15:25
Job time : 19.3 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:56:11 ; Search time 135.6 Seconds
(without alignments)
3110.164 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 51

Sequence: 1 atggcggttagaagcttctgat.....tcttgattggaccggtgct 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	470	3	US-09-731-872-24 Sequence 24, Appl
2	51	100.0	470	3	US-09-876-997-24 Sequence 24, Appl
3	51	100.0	470	9	US-10-643-836-24 Sequence 24, Appl
4	51	100.0	691	3	US-09-731-872-19 Sequence 19, Appl
5	51	100.0	691	3	US-09-876-997-19 Sequence 19, Appl
6	51	100.0	691	9	US-10-643-836-19 Sequence 19, Appl
7	51	100.0	1337	8	US-10-467-046-2 Sequence 2, Appl
8	51	100.0	1480	7	US-10-467-535-16 Sequence 16, Appl
9	51	100.0	1492	6	US-10-104-047-1704 Sequence 1704, Ap
10	51	100.0	1898	5	US-10-050-704-27 Sequence 27, Appl
11	51	100.0	1898	7	US-10-798-512-27 Sequence 27, Appl
12	51	100.0	1898	9	US-10-472-533-175 Sequence 175, App
13	51	100.0	10115	8	US-10-467-046-1 Sequence 1, Appl
14	26	51.0	526	4	US-09-925-065A-21218 Sequence 21218, A
15	25.6	50.2	74648	8	US-10-719-993-6875 Sequence 6875, Ap
16	25.6	50.2	115434	8	US-10-719-993-7058 Sequence 7058, Ap
17	25	49.0	534	8	US-10-653-047-858 Sequence 858, App
18	24.2	47.5	1179	6	US-10-243-552-110 Sequence 110, App
19	24.2	47.5	4156	6	US-10-104-047-317 Sequence 317, App
20	23.4	45.9	4494	6	US-10-377-614-3 Sequence 3, Appl
21	23.4	45.9	17758	6	US-10-377-614-7 Sequence 7, Appl
22	23.2	45.5	573	3	US-09-738-973-468 Sequence 468, App
23	23.2	45.5	573	3	US-09-854-133-468 Sequence 468, App

ALIGNMENTS

RESULT 1

US-09-731-872-24
; Sequence 24, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLAVLIWTGA/EN
US-09-731-872-24

Query Match 100.0%; Score 51; DB 3; Length 470;

Best Local Similarity 100.0%; Pred. No. 2.7e+10;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTAGAGTCTTGATGCTCTCGTCTTGTGATTGGACCGGTGCT 51

Db 137 ATGGCGGTAGAGTCTTGATGCTCTCGTCTTGTGATTGGACCGGTGCT 187

RESULT 2

US-09-876-997-24

; Sequence 24, Application US/09876997

; Publication No. US20030152921A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLLAVLIWTGA/EN
US-09-876-997-24

Query Match 100.0%; Score 51; DB 3; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 51
|||||
Db 137 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 187

RESULT 3

US-10-643-836-24
; Sequence 24, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLLAVLIWTGA/EN
US-10-643-836-24

Query Match 100.0%; Score 51; DB 9; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 51
|||||
Db 137 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 187

RESULT 4

US-09-731-872-19
; Sequence 19, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLLAVLIWTGA/EN
US-09-731-872-19

Query Match 100.0%; Score 51; DB 3; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 51
|||||
Db 42 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 92

RESULT 5

US-09-876-997-19
; Sequence 19, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-09-876-997-19

Query Match      100.0%; Score 51; DB 3; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCT 51
    |||||||
Db 42 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCT 92
    |||||||

RESULT 6
US-10-643-836-19
; Sequence 19, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Joberet, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-10-643-836-19

Query Match      100.0%; Score 51; DB 9; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCT 51
    |||||||
Db 42 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCT 92
    |||||||

RESULT 7
US-10-467-046-2
; Sequence 2, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
```

```
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)..(103)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (104)..(526)
US-10-467-046-2

Query Match      100.0%; Score 51; DB 8; Length 1337;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCT 51
    |||||||
Db 53 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCT 103
    |||||||

RESULT 8
US-10-467-535-16
; Sequence 16, Application US/10467535
; Publication No. US20040146970A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; YAO, Monique G.
; APPLICANT: ISON, Craig H.; LU, Yan
; APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
; APPLICANT: BAUGHN, Mariah R.; DING, Li
; APPLICANT: XU, Yuming; GIETZEN, Kimberly J.
; APPLICANT: TANG, Tom Y.; LAL, Preeti G.
; APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
; APPLICANT: LU, Dyrung Aina M.; RICHARDSON, Thomas W.
; APPLICANT: TRAN, Oyen K.; KHARE, Reena
; APPLICANT: CHAWLA, Narinder K.
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-0903 USN
; CURRENT APPLICATION NUMBER: US/10/467,535
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/US02/03715
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/268,111
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/271,175
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,503
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/274,552
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 1480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 5284076CB1
US-10-467-535-16

Query Match      100.0%; Score 51; DB 7; Length 1480;
Best Local Similarity 100.0%; Pred. No. 3.5e-10;
```

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
DB 29 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 79

RESULT 9

US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1704

Query Match 100.0%; Score 51; DB 6; Length 1492;

Best Local Similarity 100.0%; Pred. No. 3.5e-10; Length 1492;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
DB 52 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 102

RESULT 10

US-10-050-704-27
; Sequence 27, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-050-704-27

Query Match 100.0%; Score 51; DB 5; Length 1898;

Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
DB 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 65

RESULT 11

US-10-798-512-27
; Sequence 27, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-798-512-27

Query Match 100.0%; Score 51; DB 7; Length 1898;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 51
DB 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCT 65

RESULT 12

US-10-472-533-175
; Sequence 175, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS906PCT
; CURRENT APPLICATION NUMBER: US/10/472,533
; CURRENT FILING DATE: 2003-09-20
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 650
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1398)..(1398)

```
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1428)..(1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-472-533-175

Query Match      100.0%; Score 51; DB 9; Length 1898;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 51
    |||||||
Db 15 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 65

RESULT 13
US-10-467-046-1
; Sequence 1, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-1020US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 10115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2387...2501
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5443...5646
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6649...6747
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8907...9774
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2986)..(2986)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3847)..(3848)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3879)..(3879)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4105)..(4105)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4107)..(4109)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature

; OTHER INFORMATION: (7347)..(7348)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8197)..(8197)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8293)..(8293)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8334)..(8336)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8342)..(8349)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9942)..(9942)
; OTHER INFORMATION: n is a, c, g, or t
US-10-467-046-1

Query Match      100.0%; Score 51; DB 8; Length 10115;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 51
    |||||||
Db 2439 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 2489

RESULT 14
US-09-925-065A-21218/c
; Sequence 21218, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21218
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-21218

Query Match      51.0%; Score 26; DB 4; Length 526;
Best Local Similarity 70.0%; Pred. No. 3.4;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 TGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 51
    |||||||
Db 131 TGTGGTTAAAGGATTCTTCTTCCTCGGTTCCTCATATGGACCTGTGCT 82

RESULT 15
US-10-719-993-6875
```

```
; Sequence 6875, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6875
; LENGTH: 74648
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6875
```

```
Query Match      50.2%; Score 25.6; DB 8; Length 74648;
Best Local Similarity 70.8%; Pred. No. 15;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      2  TGGCGTTAGAGTCTTGATGCTCCTCGCTGCTCTTGATTTGGACCGGTG 49
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      839  TGGTCTTGGGAACTTTAAGCTCCTCTCAGTTTTCAGATGGCCAGGTG 886
```

Search completed: May 9, 2006, 15:02:49
Job time : 137.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 13:13:06 ; Search time 105 Seconds
(without alignments)
1978.089 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 51

Sequence: 1 atggcgtagaagcttgat.....tcttgattggaccggtgct 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_New.*
1: /SIDSS/ptodata/1/pubpna/US08_NEW_PUB.seq1.*
2: /SIDSS/ptodata/1/pubpna/US06_NEW_PUB.seq1.*
3: /SIDSS/ptodata/1/pubpna/US07_NEW_PUB.seq1.*
4: /SIDSS/ptodata/1/pubpna/US08_NEW_PUB.seq1.*
5: /SIDSS/ptodata/1/pubpna/PCT_NEW_PUB.seq1.*
6: /SIDSS/ptodata/1/pubpna/US03_NEW_PUB.seq1.*
7: /SIDSS/ptodata/1/pubpna/US09_NEW_PUB.seq1.*
8: /SIDSS/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
9: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq1.*
10: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq2.*
11: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq3.*
12: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq4.*
13: /SIDSS/ptodata/1/pubpna/US10_NEW_PUB.seq5.*
14: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq1.*
15: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
16: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
17: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq4.*
18: /SIDSS/ptodata/1/pubpna/US11_NEW_PUB.seq5.*
19: /SIDSS/ptodata/1/pubpna/US60_NEW_PUB.seq1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	497	11	US-10-475-075-564
2	51	100.0	500	11	US-10-475-075-564
3	51	100.0	1492	18	US-11-072-512-1704
4	26	51.0	526	7	US-09-925-065A-21218
5	26	51.0	526	11	US-10-301-480-1122455
6	26	51.0	526	12	US-10-301-480-735864
7	24.2	47.5	4156	9	US-10-511-937-453
8	24.2	47.5	4156	18	US-11-072-512-317
9	23.6	46.3	34554	11	US-10-893-483-49
10	22.4	43.9	741	12	US-10-301-480-552778
11	22.4	43.9	741	12	US-10-301-480-1166187
12	22.4	43.9	128361	9	US-10-505-928-151
13	22.2	43.5	625	7	US-09-925-065A-456973
14	22.2	43.5	635	12	US-10-301-480-514932

c 15	22.2	43.5	635	12	US-10-301-480-1128341	Sequence 1128341,
c 16	22.2	43.5	1088	18	US-11-188-298-106	Sequence 106, App
c 17	22	43.1	572	7	US-09-925-065A-806364	Sequence 806364,
18	22	43.1	573	7	US-09-925-065A-742070	Sequence 742070,
19	22	43.1	587	12	US-10-301-480-375274	Sequence 375274,
20	22	43.1	587	12	US-10-301-480-988683	Sequence 988683,
c 21	22	43.1	588	12	US-10-301-480-375273	Sequence 375273,
c 22	22	43.1	591	7	US-09-925-065A-298636	Sequence 298636,
c 23	22	43.1	591	7	US-09-925-065A-298637	Sequence 298637,
c 24	22	43.1	591	7	US-09-925-065A-298638	Sequence 298638,
c 25	22	43.1	591	12	US-10-301-480-375275	Sequence 375275,
c 26	22	43.1	591	12	US-10-301-480-988682	Sequence 988682,
c 27	22	43.1	591	12	US-10-301-480-988684	Sequence 988684,
c 28	22	43.1	921	10	US-10-750-185-64489	Sequence 64489, A
c 29	22	43.1	921	10	US-10-750-623-64489	Sequence 637082,
c 30	21.8	42.7	533	7	US-09-925-065A-637082	Sequence 637082,
c 31	21.8	42.7	537	11	US-10-301-480-49984	Sequence 49984, A
c 32	21.8	42.7	537	12	US-10-301-480-663393	Sequence 663393,
c 33	21.8	42.7	625	7	US-09-925-065A-71644	Sequence 71644, A
c 34	21.8	42.7	625	7	US-09-925-065A-71645	Sequence 71645, A
c 35	21.8	42.7	625	11	US-10-301-480-172883	Sequence 172883,
c 36	21.8	42.7	625	11	US-10-301-480-172884	Sequence 172884,
c 37	21.8	42.7	625	12	US-10-301-480-786292	Sequence 786292,
c 38	21.8	42.7	625	12	US-10-301-480-786293	Sequence 786293,
c 39	21.8	42.7	1139	7	US-09-925-065A-75377	Sequence 75377, A
c 40	21.8	42.7	1139	11	US-10-301-480-176616	Sequence 176616,
c 41	21.8	42.7	1139	12	US-10-301-480-790025	Sequence 790025,
c 42	21.8	42.7	1310	10	US-10-750-185-50706	Sequence 50706, A
c 43	21.8	42.7	1310	10	US-10-750-623-50706	Sequence 50706, A
c 44	21.8	42.7	317876	10	US-10-995-561-13227	Sequence 13227, A
c 45	21.6	42.4	471	7	US-09-925-065A-650301	Sequence 650301,

ALIGNMENTS

RESULT 1

US-10-475-075-564
; Sequence 564, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 564
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 52..495
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 52..102
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAFLVITGA/EN
US-10-475-075-564

Query Match 100.0%; Score 51; DB 11; Length 497;

Best Local Similarity 100.0%; Pred. No. 8.3e-11;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 51
|||||
Db 52 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 102

RESULT 2'
US-10-475-075-563
; Sequence 563, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Giobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 563
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 72..500
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 72..122
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1..3
; OTHER INFORMATION: n = a, g, c or t
US-10-475-075-563

Query Match 100.0%; Score 51; DB 11; Length 500;
Best Local Similarity 100.0%; Pred. No. 8.3e-11;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 51
|||||
Db 72 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 122

RESULT 3
US-11-072-512-1704
; Sequence 1704, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYU
; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-1704

Query Match 100.0%; Score 51; DB 18; Length 1492;
Best Local Similarity 100.0%; Pred. No. 9.7e-11;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 51
|||||
Db 52 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 102

RESULT 4
US-09-925-065A-21218/C
; Sequence 21218, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21218
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-21218

Query Match 51.0%; Score 26; DB 7; Length 526;
Best Local Similarity 70.0%; Pred. No. 1.3;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2 TGGCGTTAGAGTCTTGATGCTCCTCGCTGTCCTTGATTGGACCGGTGCT 51
|||||
Db 131 TGTGCTTAAAGGATTCCTCTCCGTTCTTCATATGGACCTGTGCT 82

RESULT 5
US-10-301-480-122455/c
; Sequence 122455, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137


```

; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 453
; LENGTH: 4156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-453

Query Match          47.5%;   Score 24.2; DB 9;   Length 4156;
Best Local Similarity 78.4%;   Pred. No. 9.7;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      15  CTTGATGCTCCTCGCTGCTCTTGATTTGGACCGGTGCT 51
          |||||  |||||  |||||  |||||  |||||  |||||
Db      1030 CTTGATGCTTGTCACTGTCATGATGTCGCCCTTGCT 1066

RESULT 8
US-11-072-512-317
; Sequence 317, Application US/11072512
; Publication No. US2006029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cdNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 317
; LENGTH: 4156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-317

Query Match          47.5%;   Score 24.2; DB 18;   Length 4156;
Best Local Similarity 78.4%;   Pred. No. 9.7;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      15  CTTGATGCTCCTCGCTGCTCTTGATTTGGACCGGTGCT 51
          |||||  |||||  |||||  |||||  |||||  |||||
Db      1030 CTTGATGCTTGTCACTGTCATGATGTCGCCCTTGCT 1066

RESULT 9
US-10-893-483-49/c

```

```
; Sequence 49, Application US/10893483
; Publication No. US20060026696A1
; GENERAL INFORMATION:
; APPLICANT: Buelow, Roland
; APPLICANT: Platzner, Josef
; APPLICANT: Schooten, Wim van
; TITLE OF INVENTION: Humanized Immunoglobulin Loci
; FILE REFERENCE: 39691-0007A
; CURRENT APPLICATION NUMBER: US/10/893,483
; PRIOR FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: 60/487,733
; PRIOR FILING DATE: 2003-07-15
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 34554
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-10-893-483-49

Query Match 46.3%; Score 23.6; DB 11; Length 34554;
Best Local Similarity 76.3%; Pred. No. 23;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 14 TCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCT 51
||||| ||||| ||||| ||||| ||||| |||||
Db 6906 TCTTCAGGCTCTCTTTGCTTTATTGGCGTGGTGT 6869

RESULT 10
US-10-301-480-552778
; Sequence 552778, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 552778
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-552778

Query Match 43.9%; Score 22.4; DB 12; Length 741;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 TGGCGTTAGAACTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTG 49
||||| ||||| ||||| ||||| ||||| |||||
Db 230 TGGTATTAGTAGTCTCCAAAGGTCCTCACTGTTGGCTGTAGAGTG 277

RESULT 11
US-10-301-480-1166187
; Sequence 1166187, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1166187
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1166187

Query Match 43.9%; Score 22.4; DB 12; Length 741;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 TGGCGTTAGAACTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTG 49
||||| ||||| ||||| ||||| ||||| |||||
Db 230 TGGTATTAGTAGTCTCCAAAGGTCCTCACTGTTGGCTGTAGAGTG 277

RESULT 12
US-10-505-928-151/c
; Sequence 151, Application US/10505928
; Publication No. US20060089532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 151
; LENGTH: 128361
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-151

Query Match 43.9%; Score 22.4; DB 9; Length 128361;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 TGGCGTTAGAACTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTG 49
||||| ||||| ||||| ||||| ||||| |||||
Db 24897 TGGCTTTGACATCTCTGATTTCTCCCACTGATTGCTTGGAAACCAGGG 24850

RESULT 13
US-09-025-065A-456973/c
; Sequence 456973, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 456973
```

```

; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-456973

QY 8 TAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTGG 42
Db 353 TAGAGGTCTTGAAGCTCCTCCCTCTGTAGAAGTGG 319
```

Query Match 43.5%; Score 22.2; DB 7; Length 625;
Best Local Similarity 77.1%; Pred. No. 49;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 8 TAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTGG 42
Db 353 TAGAGGTCTTGAAGCTCCTCCCTCTGTAGAAGTGG 319
```

RESULT 14

```

US-10-301-480-514932/c
; Sequence 514932, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 514932
; LENGTH: 635
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-514932
```

Query Match 43.5%; Score 22.2; DB 12; Length 635;
Best Local Similarity 77.1%; Pred. No. 49;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 8 TAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTGG 42
Db 353 TAGAGGTCTTGAAGCTCCTCCCTCTGTAGAAGTGG 319
```

RESULT 15

```

US-10-301-480-1128341/c
; Sequence 1128341, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1128341
; LENGTH: 635
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1128341
```

Query Match 43.5%; Score 22.2; DB 12; Length 635;
Best Local Similarity 77.1%; Pred. No. 49;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 8 TAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTGG 42
Db 353 TAGAGGTCTTGAAGCTCCTCCCTCTGTAGAAGTGG 319
```

Search completed: May 9, 2006, 13:33:05
Job time : 106 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:54:43 ; Search time 630 Seconds
(without alignments)
3787.526 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 51

Sequence: 1 atggcgttagaagcttgat.....tcttgattggaccggtgct 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *
10: gb_est10: *
11: gb_est11: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	51	100.0	331	1	AI138881 qd99c04.x
2	51	100.0	516	2	BE972717
3	51	100.0	626	3	BI463012
4	51	100.0	753	5	BU561351
5	51	100.0	776	2	EG772527
6	51	100.0	778	3	BI459538
7	51	100.0	878	3	BI520154
8	38	74.5	577	5	BI118111
9	33.4	65.5	598	2	EG081002
10	31.8	62.4	629	5	BY731676
11	31.8	62.4	862	7	CO810251
12	27.8	54.5	834	9	AQ743481
13	26.8	52.5	859	9	CC815351
14	26.6	52.2	621	10	AG153276
15	26.4	51.8	738	6	CF811959
16	26.4	51.8	813	7	CO029792
17	26.4	51.8	902	6	CF819479
18	26.4	51.8	961	6	CF816222
19	25.8	50.6	731	2	BI225701
20	25.4	49.8	205	1	BB008763
21	25.4	49.8	324	1	BB130781
22	25.4	49.8	816	10	AG483623

23	25.2	49.4	804	4	CNSOGX06
24	25.2	49.4	827	4	CNSOGJ6V
C 25	24.8	48.6	581	10	CW093458
C 26	24.8	48.6	1088	10	CZ362098
27	24.6	48.2	429	5	BY636904
C 28	24.6	48.2	730	7	CO960964
29	24.4	47.8	129	10	CW144937
30	24.4	47.8	385	10	CW516351
31	24.4	47.8	536	10	CW533066
C 32	24.4	47.8	650	9	AZ572445
C 33	24.4	47.8	657	6	CF481997
34	24.4	47.8	701	7	CN242154
35	24.2	47.5	199	7	CN370178
C 36	24.2	47.5	307	1	AA707067
37	24.2	47.5	507	5	EX283182
38	24.2	47.5	542	2	BE265311
39	24.2	47.5	588	7	CN370177
40	24.2	47.5	647	3	BM042024
41	24.2	47.5	666	7	CN370181
42	24.2	47.5	682	7	CR980458
43	24.2	47.5	694	8	DN999224
44	24.2	47.5	754	2	BG750320
45	24.2	47.5	769	6	CB989869

ALIGNMENTS

RESULT 1
LOCUS AI138881/C
DEFINITION qd99c04.xl Soares_testis_NHT Homo sapiens cdna clone IMAGE:1737606
3', mRNA sequence.
ACCESSION AI138881
VERSION AI138881.1 GI:3644853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 331)
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 395 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES

Location/Qualifiers
1..331
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1737606"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares_testis_NHT"
note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cdna was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5'.
TGTTCACCATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cdna was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT773 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 51; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 331 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 281
|||||

RESULT 2

BE972717 516 bp mRNA linear EST 04-OCT-2000
LOCUS 601652136F1 NTH_MGC_82 Homo sapiens cDNA clone IMAGE:3935288 5',
DEFINITION mRNA sequence.

ACCESSION BE972717

VERSION BE972717

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM777 row: h column: 09

High quality sequence stop: 505.

FEATURES

source

1..516

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3935288"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_82"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:

SfiI (ggccgctggcc); Site 2: SfiI (ggccattggcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:

5'-ATTCTAGCGCGAGCGCGCATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size

1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained

inserts by PCR. This library was enriched for full-length

clones and was constructed by Clontech Laboratories (Palo

Alto, CA)."

ORIGIN

Query Match 100.0%; Score 51; DB 2; Length 516;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 120 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 170
|||||

RESULT 3

BI463012

LOCUS

DEFINITION

603204535F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5270384 5',
mRNA sequence.

ACCESSION BI463012

VERSION BI463012.1

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 626)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11682 row: e column: 09

High quality sequence stop: 624.

FEATURES

source

1..626

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5270384"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/note="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.2 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity

Matches

51; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51

|||||

Db

57 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 107

|||||

RESULT 4

BUS61351/c

LOCUS

DEFINITION

BUS61351

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 753)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Query Match

Best Local Similarity

Matches

51; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51

|||||

Db

120 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 170

|||||

RESULT 3

BUS61351

LOCUS

DEFINITION

BUS61351

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 753)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium information can be
http://image.llnl.gov
Plate: LLCM2815 row: g column: 12
High quality sequence stop: 442.
Location/Qualifiers
1. .753
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6592356"
/lab host="DH10B (T1 phage-resistant)"
/clone lib="NIH MGC 82"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccgctcgcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGCGCGCACATG-dt (30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."

FEATURES

source

ORIGIN

Query Match 100.0%; Score 51; DB 2; Length 776;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTCTGATTGGACCGGTGCT 51

Db 57 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTTCTGATTGGACCGGTGCT 107

RESULT 6

BI459538 778 bp mRNA linear EST 21-AUG-2001
LOCUS 603200545F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5266642 5',
DEFINITION mRNA sequence.

ACCESSION BI459538

VERSION BI459538.1 GI:15250194

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1 (bases 1 to 778)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rc@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1672 row: i column: 11

High quality sequence stop: 674.

FEATURES

source

Location/Qualifiers
1. .778

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5266642"

/lab host="DH10B"

/clone lib="NIH MGC 97"

/notes="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gtcgag); Oligo-dt primed using primer

5'-TTTTTTTTTTTTTNN-3', size-selected for average

insert size 2.2 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

FEATURES

source

Location/Qualifiers
1. .776

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5266642"

/lab host="DH10B"

/clone lib="NIH MGC 97"

/notes="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gtcgag); Oligo-dt primed using primer

5'-TTTTTTTTTTTTTNN-3', size-selected for average

insert size 2.2 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 51; DB 3; Length 778;
 Best Local Similarity 100.0%; Pred. No. 2.7e-08;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTGACCGGTGCT 51
 |||||
 Db 191 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTGACCGGTGCT 241
 |||||

RESULT 7

B1520154
 LOCUS 603071295F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163455 5',
 DEFINITION mRNA sequence.

ACCESSION

B1520154

VERSION

B1520154

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE

1 (bases 1 to 878)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11405 row: m column: 24
 High quality sequence stop: 851.

FEATURES

source

1..878
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5163455"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /clone_lib="NIH MGC_119"
 /note="Organ: brain; Vector: pCMV-SF0RT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 51; DB 3; Length 878;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTGACCGGTGCT 51
 |||||

Db

30 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTGACCGGTGCT 80
 |||||

RESULT 8

BX118111

LOCUS

DEFINITION BX118111 Soares_testis NHT Homo sapiens cDNA clone IMAGE:1030619, mRNA sequence.

ACCESSION

VERSION BX118111.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE

1 (bases 1 to 577)

AUTHORS

Radelof, U., Schneider, D. and Korn, B.

TITLE

Human Unigeneset - RZPD3

JOURNAL

Unpublished (2003)

COMMENT

Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGp998D122574.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Human Unigeneset - RZPD3 (RZPDLIB No.972)
 http://www.rzpd.de/CloneCards/cgi-

bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

M13r, Primer sequence: TTTCACAGGAACAGCTATGAC.

FEATURES

source

1..577
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGp998D122574 ; IMAGE:1030619"
 /sex="male"
 /lab_host="DH10B"
 /clone_lib="Soares testis NHT"
 /note="Vector: pT7T3-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5',
 TGTTACCAATCTGAAGTGGGCGCGCCCAATTTTTTTTTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 74.5%; Score 38; DB 5; Length 577;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TCTTGATGCTCTCGTGTCTTGATTGGACCGGTGCT 51
 |||||

Db 1 TCTTGATGCTCTCGTGTCTTGATTGGACCGGTGCT 38
 |||||

RESULT 9

BG081002

LOCUS

DEFINITION BG081002 H3059H12-5 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3059H12 5', mRNA sequence.

ACCESSION

VERSION BG081002

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	Mus musculus	ORGANISM	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus. 1 (bases 1 to 629)		
TITLE	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tonita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszewski, B., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.		
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
PUBLISHED	Nature 420, 563-573 (2002)		
COMMENT	12466851		
	Contact: Yoshihide Hayashizaki		
	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute		
	The Institute of Physical and Chemical Research (RIKEN)		
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan		
	Tel: 81-45-503-9222		
	Fax: 81-45-503-9216		
	Email: genome-resesc.riken.jp, URL: http://genome.gsc.riken.jp/		
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.		
	Direct Submission		
	Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)		
	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)		
	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)		
	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)		
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.		
	Please visit our web site (http://genome.gsc.riken.go.jp) for further details.		
FEATURES	Location/Qualifiers		
source	1..598		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/db_xref="Ensembl:ENST:H3059H12-5"		
	/clone="H3059H12"		
	/sex="Clones arrayed from a variety of cDNA libraries"		
	/dev_stage="Clones arrayed from a variety of cDNA libraries"		
	/lab_host="DH10B"		
	/clone_lib="NIA Mouse 15K cDNA Clone Set"		
	/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryos cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."		
ORIGIN			
Query Match	55.5%; Score 33.4; DB 2; Length 598;		
Best Local Similarity	78.4%; Pred. No. 0.11;		
Matches	40; Conservative 0; Mismatches 11; Indels 0; Gaps 0;		
QY	1 ATGGCGCTTAGAGCTTGTATGCTCCGCTGCTTCATTTGGACCGGTCT 51		
	41 ATGGCGCTTAGAGCTTGTATGCTTCGCTGCTTCGCTTCGCTGCTGT 91		
Db			
RESULT 10			
BY731676			
LOCUS	629 bp mRNA linear EST 17-DEC-2002		
DEFINITION	BY731676 RIKEN full-length enriched, 2 days pregnant adult female		
ACCESSION	ovary Mus musculus cDNA clone E330031A07 5', mRNA sequence.		
VERSION	BY731676		
KEYWORDS	BY731676.1 GI:27144803		
SOURCE	Mus musculus (house mouse)		

Query Match	54.5%; Score 27.8; DB 9; Length 834;	
Best Local Similarity	74.5%; Pred. No. 15;	
Matches	35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;	
Qy	2 TGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGT 48 	
Db	231 TGGCGCGGATGCTTTGATGATCCTGCTGCTGGGGATTTGCATCAGT 277	
RESULT 13		
CC815351/c		
LOCUS	CC815351	
DEFINITION	ZMMBBC0521C24r ZMMBbc Zea mays genomic clone ZMMBBC0521C24 3', genomic survey sequence.	
ACCESSION	CC815351	
VERSION	CC815351.1 GI:32821976	
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
AUTHORS	Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Rouzard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.	
TITLE	Sequencing of the maize genome at PGIR (2003b)	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Bharti.A.K. Dr.Joachim Messing's lab The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University 190 Frelinghuysen Road, Piscataway, NJ 08854, USA Tel: 732 445 3801 Fax: 732 445 5735 Email: bharti@waksman.rutgers.edu Seq primer: SP6 Class: BAC ends High quality sequence start: 455. Location/Qualifiers 1. .859 /organism="Zea mays" /mol_type="genomic DNA" /cultivar="B73" /db_xref="taxon:4577" /clone="ZMMBBC0521C24" /lab_host="E. coli DH10B" /clone_lib="ZMMBBC" /notes="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"	
FEATURES	source	
ORIGIN		
Query Match	52.5%; Score 26.8; DB 9; Length 859;	
Best Local Similarity	73.9%; Pred. No. 36;	
Matches	34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;	
Qy	3 GGCCTTAGAGTCTTGATGCTCTCGCTGCTGCTTGTGATTTGGACCGGT 48 	
Db	139 GGCCTTAGAATACTTGCTCATCTCTCTGCTTGTGATCAGTACCTGT 94	
RESULT 14		
AG153276		
LOCUS	AG153276	
DEFINITION	Pan troglodytes DNA, clone: RP43-016P02.TU, genomic survey sequence.	
ACCESSION	AG153276	
VERSION	AG153276.1 GI:16682954	
KEYWORDS	GSS.	
SOURCE	Pan troglodytes (chimpanzee)	
ORGANISM	Pan troglodytes	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan. 1	
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.	
TITLE	BAC end sequences of Library RPCI-43	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 621) Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y. Direct Submission Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpansegsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) Clones are derived from the chimpanzee BAC library RPCI-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors. PRIMERS	
COMMENT	Sequencing: TJ	
LIBRARY	Vector : pBACe3.6 R.Site 1 : EcoRI R.Site 2 : EcoRI. Location/Qualifiers 1. .621 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" /clone="RP43-016P02.TU" /sex="male" /cell_type="lymphocytes" /clone_lib="RPCI-43 Chimpanzee Male BAC Library"	
ORIGIN		
Query Match	52.2%; Score 26.6; DB 10; Length 621;	
Best Local Similarity	76.2%; Pred. No. 41;	
Matches	32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	
Qy	2 TGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGA 43 	
Db	254 TGGNGTTCATGCTCTTGATGATCTCGCTGGCGGATTTGCA 295	
RESULT 15		
CF811959/c		
LOCUS	CF811959	
DEFINITION	EST689341 Coccidioides posadasii saprobic phase cDNA library, greater than 4kb Coccidioides posadasii cDNA clone CIBAL42 5' end, mRNA sequence.	
ACCESSION	CF811959	
VERSION	CF811959.1 GI:45917837	
KEYWORDS	EST.	
SOURCE	Coccidioides posadasii	
ORGANISM	Coccidioides posadasii Eukaryota; Fungi; Ascomycota; Peizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides. 1 (bases 1 to 738) Gardner,M.J. and Cole,G.T. Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags Unpublished (2003) Other ESTs: EST689340 Contact: Gardner MJ The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301 838 3519 Fax: 301 838 0208 Email: gardner@tigr.org Seq primer: M13 Reverse Location/Qualifiers 1. .738 /organism="Coccidioides posadasii" /mol_type="mRNA" /strain="C735"	

```
/db_xref="taxon:199306"  
/clone="CIBA142"  
/dev_stage="saprobic phase (mycelia)"  
/lab_host="E. coli DH10B, T1 phage resistant"  
/clone_lib="Coccidioides posadasii saprobic phase cDNA  
library, greater than 4kb"  
/note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;  
Coccidioides posadasii saprobic phase cDNA library, size  
fractionated cDNA > 4 kb"
```

ORIGIN

```
Query Match      51.8%; Score 26.4; DB 6; Length 738;  
Best Local Similarity 75.0%; Pred. No. 50;  
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      8 TAGAAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGTGCT 51  
        ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db      687 TAGATGTCCTGGAGGCGCCTCGCTTTTGGATTTCGAAAGGTGAT 644
```

Search completed: May 9, 2006, 14:40:01
Job time : 635 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus.n2p model

Run on: May 9, 2006, 10:01:14 ; Search time 2.02 Seconds

(without alignments)
3327.968 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 92

Sequence: 1 atggcgtagaagtcttgat.....tcttgattggaccggtgct 51

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+.n2p.model -DEV=xlh
-Q-/abs/ABSWEB/spool/US10664025/runat_08052006_173448_28291/app_query.fasta 1
-DB=A_Geneseq -Qfmt=faстан -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs07
-USER=US10664025 -CGN_1_1_241 -runat_08052006_173448_28291 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	88.0	106	4 AAG89145	Aag89145 Human sec
2	81	88.0	143	6 ABP76237	Abp76237 Human GEN
3	81	88.0	148	6 ABP76238	Abp76238 Human GEN
4	81	88.0	158	4 AAG89140	Aag89140 Human sec
5	81	88.0	158	5 ABG97353	Abg97353 Human CGD
6	81	88.0	158	5 ABG31324	Abg31324 Human GSS
7	81	88.0	158	6 ABR47931	Abp47931 Human sec
8	81	88.0	158	6 ABR00175	Abp00175 Human gen
9	81	88.0	158	7 ADB91675	Adb91675 Human sec

10	81	88.0	158	7 ADB65520	Adb65520 Human pro
11	81	88.0	158	7 ADC74338	Adc74338 Human sec
12	81	88.0	159	3 AAB38337	Aab38337 Human sec
13	47	51.1	250	2 AAW14446	Aaw14446 Carz gene
C 14	47	51.1	624	5 ABG79696	Abg79696 Tumour in
C 15	47	51.1	633	7 ADD14149	Add14149 Human sec
C 16	47	51.1	633	7 ABM85795	Abm85795 Human pro
17	47	51.1	714	8 ADX79030	Adx79030 Plant ful
18	47	51.1	996	6 ABU22513	Abu22513 Protein e
19	46	50.0	80	9 ADX40751	Adx40751 HIV Vpu p
C 20	46	50.0	101	9 ADX40372	Adx40372 HIV Tat p
21	46	50.0	102	8 ADY24002	Ady24002 Plant ful
22	46	50.0	138	8 ADY11096	Ady11096 Plant ful
C 23	46	50.0	209	7 ABO74477	Abp74477 Pseudomon
24	46	50.0	226	3 AAB51908	Aab51908 Gene 30 h
C 25	46	50.0	342	7 ABO74508	Abp74508 Pseudomon
26	46	50.0	345	8 ADM25498	Adm25498 Bacterial
27	46	50.0	709	8 ADT55713	Adt55713 Plant pol
28	46	50.0	862	7 ADC99860	Adc99860 Murine LD
29	46	50.0	862	8 ADI27189	Adi27189 Mouse LRP
30	46	50.0	862	8 ADI27190	Adi27190 Mouse LRP
31	46	50.0	864	8 ADI27191	Adi27191 Mouse LRP
C 32	45	48.9	69	3 AAG47936	Aag47936 Arabidops
C 33	45	48.9	70	3 AAG09569	Aag09569 Arabidops
C 34	45	48.9	70	3 AAG47945	Aag47945 Arabidops
35	45	48.9	80	9 ADX40648	Adx40648 HIV Vpu p
36	45	48.9	81	9 ADX40735	Adx40735 HIV Vpu p
37	45	48.9	104	5 ABB94049	Abb94049 Human sec
38	45	48.9	104	5 ABB94025	Abb94025 Human sec
39	45	48.9	104	5 ABG65110	Abg65110 Human alb
40	45	48.9	104	5 ABG65111	Abg65111 Human alb
41	45	48.9	104	8 ADL78378	Adl78378 Albumin f
42	45	48.9	104	8 ADL78377	Adl78377 Albumin f
43	45	48.9	120	4 AAU20830	Aau20830 Human nov
44	45	48.9	323	8 ADU06834	Adu06834 Cell adhe
45	45	48.9	335	4 AAG72080	Aag72080 Human olf

ALIGNMENTS

RESULT 1

AAG89145

ID AAG89145 standard; protein; 106 AA.

XX AAG89145;

XX AC AAG89145;

DT 11-SEP-2001 (first entry)

XX DE Human secreted protein, SEQ ID NO: 265.

XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

XX KW GENSET.

XX OS Homo sapiens.

XX PN WO200142451-A2.

XX PD 14-JUN-2001.

XX PF 07-DEC-2000; 2000WO-IB001938.

XX PR 08-DEC-1999; 99US-0169629P.

XX PR 06-MAR-2000; 2000US-0187470P.

XX (GEST) GENSET.

PI Dunas Milne Edwards J, Bougueleret L, Jobert S;

DR WPI; 2001-367870/38.

XX N-PSDB; AAH64748.

PT Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of

PT diseases, and for diagnosis of those diseases.

PS Claim 21; Page 806; 921pp; English.

XX
CC The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention

XX
SQ Sequence 106 AA;

Alignment Scores:
Pred. No.: 0.00017 Length: 106
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x AAG89145 (1-106)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGTGTCTTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla 17

RESULT 2

ABP76237

ID ABP76237 standard; protein; 143 AA.

XX

AC ABP76237;

XX 21-FEB-2003 (first entry)

DT Human GENSET protein SEQ ID 787.

DE

XX

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;

KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;

KW inflammatory disease; immune disorder; neuromuscular; toxicity;

KW central nervous system; cardiovascular; gastrointestinal.

XX

OS Homo sapiens.

XX

PN WO200283898-A1.

XX

PD 24-OCT-2002.

XX

PF 18-APR-2001; 2001WO-IB000914.

XX

PR 18-APR-2001; 2001WO-IB000914.

XX

PA (GEST) GENSET.

XX

PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;

XX

DR WPI; 2003-075548/07.

XX

XX New GENSET polynucleotides and polypeptides, useful for treating heavy

PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the

PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the

PT toxicity.

PS Claim 14; Page 680; 735pp; English.

XX
CC The present invention relates to novel GENSET polynucleotides (ABZ36404-ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides and polypeptides are useful in screening and diagnostic assays for abnormal GENSET expression and/or biological activity. They are also useful for screening of compounds for treating or preventing GENSET-related disorders, such as heavy metal toxicity, cancer, inflammatory diseases, immune disorders, and the neuromuscular, central nervous system (CNS), cardiovascular or gastrointestinal effects of the toxicity

XX
SQ Sequence 143 AA;

Alignment Scores:
Pred. No.: 0.000172 Length: 143
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x ABP76237 (1-143)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGTGTCTTGATTGACCGGTGCT 51
|||||

Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla 17

RESULT 3

ABP76238

ID ABP76238 standard; protein; 148 AA.

XX

AC ABP76238;

XX

DT 21-FEB-2003 (first entry)

XX Human GENSET protein SEQ ID 788.

DE

XX

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;

KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;

KW inflammatory disease; immune disorder; neuromuscular; toxicity;

KW central nervous system; cardiovascular; gastrointestinal.

XX

OS Homo sapiens.

XX

PN WO200283898-A1.

XX

PD 24-OCT-2002.

XX

PF 18-APR-2001; 2001WO-IB000914.

XX

PR 18-APR-2001; 2001WO-IB000914.

XX

PA (GEST) GENSET.

XX

PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;

XX

DR WPI; 2003-075548/07.

XX

XX New GENSET polynucleotides and polypeptides, useful for treating heavy

PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the

PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the

PT toxicity.

XX

PS Claim 14; Page 680; 735pp; English.

XX

XX The present invention relates to novel GENSET polynucleotides (ABZ36404-ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides

CC and polypeptides are useful in screening and diagnostic assays for

CC abnormal GENSET expression and/or biological activity. They are also

CC useful for screening of compounds for treating or preventing GENSET-

CC related disorders, such as heavy metal toxicity, cancer, inflammatory

CC diseases, immune disorders, and the neuromuscular, central nervous system
 CC (CNS), cardiovascular or gastrointestinal effects of the toxicity

XX
 SQ Sequence 148 AA;

Alignment Scores:
 Pred. No.: 0.000172 Length: 148
 Score: 81.00 Matches: 17
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 88.0% Indels: 0
 DB: 6 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x ABP76238 (1-148)

Qy 1 ATGCGGTTAGAGCTTGTATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 51
 |||||
 Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 4

ABG89140
 ID AAG89140 standard; protein; 158 AA.

XX
 AC AAG89140;

XX
 DT 11-SEP-2001 (first entry)

XX
 DE Human secreted protein, SEQ ID NO: 260.

XX
 KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KW GENSET.

XX
 OS Homo sapiens.

XX
 PN WO200142451-A2.

XX
 PD 14-JUN-2001.

XX
 PF 07-DEC-2000; 2000WO-IB001938.

XX
 PR 08-DEC-1999; 99US-0169629P.

XX
 PR 06-MAR-2000; 2000US-0187470P.

XX
 PA (GENSET) GENSET.

XX
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX
 DR WPI; 2001-367870/38.

XX
 DR N-PSDB; AAH64743.

XX Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases.

XX
 PS Claim 21; Page 804; 921pp; English.

XX The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased GENSET
 CC gene expression by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of GENSET or by supplementing the
 CC patient's own production of GENSET polypeptides. Conversely, antisense
 CC nucleic acid molecules may be administered to down regulate GENSET
 CC expression by binding with the cells' own genes and preventing their
 CC expression. The sense and antisense nucleic acids may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and hence to determine which
 CC patients may be in need of restorative therapy. The GENSET polypeptides
 CC may be used as antigens in the production of antibodies and in assays to
 CC identify modulators (agonists and antagonists) of GENSET polypeptide
 CC expression and activity. The present sequence is a GENSET polypeptide of

CC the invention
 XX
 SQ Sequence 158 AA;

Alignment Scores:
 Pred. No.: 0.000172 Length: 158
 Score: 81.00 Matches: 17
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 88.0% Indels: 0
 DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x AAG89140 (1-158)

Qy 1 ATGCGGTTAGAGCTTGTATGCTCTCGCTGCTTGTGATTTGGACCGGTGCT 51
 |||||
 Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 5

ABG97353

ID ABG97353 standard; protein; 158 AA.

XX
 AC ABG97353;

XX
 DT 16-DEC-2002 (first entry)

XX
 DE Human CGD4, INCYTE 5284076CD1.

XX
 KW Human; cell growth; differentiation; death; CGDD; cancer;
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
 KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; poriasis;
 KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
 KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
 KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
 KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
 KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
 KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
 KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
 KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
 KW rheumatoid arthritis.

XX
 OS Homo sapiens.

XX
 PN WO200272830-A2.

XX
 PD 19-SEP-2002.

XX
 PF 08-FEB-2002; 2002WO-US003715.

XX
 PF 09-FEB-2001; 2001US-0268111P.

XX
 PR 23-FEB-2001; 2001US-0271175P.

XX
 PR 08-MAR-2001; 2001US-0274503P.

XX
 PR 09-MAR-2001; 2001US-0274552P.

XX
 PA (INCY-) INCYTE GENOMICS INC.

XX
 PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
 PI Ding L, Xu Y, Gletzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
 PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;

XX
 DR WPI; 2002-723356/78.

XX
 DR N-PSDB; ABS78646.

XX New human proteins associated with cell growth, differentiation and
 PT death, useful for diagnosing, treating or preventing autoimmune or
 PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
 PT atherosclerosis or hepatitis.

XX
 PS Claim 1; Page 146; 181pp; English.

XX The invention relates to an isolated polypeptide comprising CGDD1-12
 CC (cell growth, differentiation and death), a naturally occurring amino
 CC acid sequence at least 90% identical to CGDD, a biologically active

CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence
CC represents a CGDD protein
XX
SQ Sequence 158 AA;

Alignment Scores:
Pred. No.: 0.000172 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x ABG97353 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTCGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 6
ABG31324
ID ABG31324 standard; protein; 158 AA.
XX
AC ABG31324;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human GSSP3 polypeptide.
XX
KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
KW body mass; serum glucose regulation; body weight loss; obesity;
KW metabolic-related disorder; impaired glucose tolerance; stroke;
KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
KW hypertension; syndrome C; type I diabetes; type II diabetes;
KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
KW lipodystrophy; physical performance; exercise; dyslexia; schizophrenia;
KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder.
XX
OS Homo sapiens.

Key Location/Qualifiers
FH 1. .17
FT Peptide /label= Signal_peptide
FT 18. 158
FT Protein /label= Mature_GSSP3_polypeptide
XX

WO200260466-A2.

XX

PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-IB001333.
XX
PR 02-FEB-2001; 2001US-0266156P.
XX
XX (GEST) GENSET.
XX
PI Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
PI Bihain B;
XX
XX WPI; 2002-608487/65.
DR N-PSDB; ABK90053.
XX
XX Reducing circulating glucose levels or increasing insulin sensitivity,
XX useful for reducing body mass or preventing body weight gain, comprises
XX administering composition comprising GSSP3 polypeptide.
PS Claim 1; Page 96-97; 97pp; English.

XX The present invention relates to the isolation of human GSSP3
CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
CC polypeptide reduces circulating blood glucose levels, increases insulin
CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
CC polynucleotide sequences are useful in serum glucose regulation, fatty
CC acid metabolism, body weight loss, and prevention of body weight gain.
CC Compositions comprising GSSP3 polypeptides are useful for controlling
CC blood glucose levels, for treating metabolic-related diseases or
CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
CC syndrome C, type I or II diabetes, diabetes related complications,
CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
CC acanthosis nigricans, leprechaunism and lipodystrophy). The polypeptides
CC are also useful to improve physical performance during work or exercise,
CC and to treat dyslexia, attention-deficit disorder, attention-
CC deficit/hyperactivity disorder, and psychiatric disorders such as
CC schizophrenia. The present sequence represents human GSSP3 polypeptide
XX

SQ Sequence 158 AA;

Alignment Scores:
Pred. No.: 0.000172 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x ABG31324 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTCGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 7
ABR47931
ID ABR47931 standard; protein; 158 AA.
XX
AC ABR47931;
XX

DT 12-JUN-2003 (first entry)

XX Human secreted protein, SEQ ID 822.

XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
KW vulnery; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder.

OS Homo sapiens.

XX WO200295010-A2.

XX

XX Human secreted protein #SEQ ID 621.
XX Secreted protein; gene therapy; antidiabetic; diabetes; human.
XX Homo sapiens.
XX WO2003004622-A2.
XX 16-JAN-2003.
XX 19-MAR-2002; 2002WO-US008124.
XX 21-MAR-2001; 2001US-0277340P.
XX 19-JUL-2001; 2001US-0306171P.
XX 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2003-229407/22.
XX Nucleic acid encoding a human secreted protein is useful in diagnosing or treating diabetes or conditions related to diabetes.
XX Claim 3; SEQ ID NO 621; 1537pp; English.
XX The invention relates to isolated nucleic acid molecules ADB91065-ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-ADB91834. Also disclosed is a recombinant vector comprising a polynucleotide of the invention, and a recombinant host cell comprising the recombinant vector. The polypeptide of the invention is useful in identifying a binding partner by contacting the polypeptide with a binding partner, and determining whether the binding partner increases or decreases activity of the polypeptide. The polypeptide, polynucleotide, antibody or its fragment, agonist or antagonist are useful for preparing a pharmaceutical composition for diagnosing or treating diabetes or conditions related to diabetes. The present sequence is that of the human immunoglobulin Fc portion used to generate fusion proteins, increasing the stability of the fused protein as compared to the secreted protein only. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 158 AA;
SQ
Alignment Scores:
Pred. No.: 0.000172 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 7 Gaps: 0
US-10-664-025-43_COPY_53_103 (1-51) x ADB91675 (1-158)
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuTrpThrGlyAla 17
RESULT 10
ADB65520
ID ADB65520 standard; protein; 158 AA.
XX ADB65520;
XX 04-DEC-2003 (first entry)
XX Human protein encoded by clone TESTI20282420.
XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration; cell regeneration; membrane protein; signal transduction-related protein;
KW

KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX
XX Homo sapiens.
XX EPI308459-A2.
XX 07-MAY-2003.
XX 28-MAR-2002; 2002EP-00007401.
XX 05-NOV-2001; 2001JP-00379298.
XX 25-JAN-2002; 2002US-00350978.
XX (HELI-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-450961/43.
XX N-PSDB; ADB63550.
XX New polynucleotides and polypeptides, useful for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy.
XX Claim 1; Page; 222pp; English.
XX The invention discloses a polynucleotide comprising a sequence selected from 1970 fully defined nucleotide sequences which encode novel polypeptides. Also claimed is a polypeptide encoded by the polynucleotide or its partial peptide, an antibody binding to the polypeptide or peptide of the polynucleotide, immunologically assaying the polypeptide or peptide of the polynucleotide by contacting the polypeptide or peptide with the antibody of the encoded protein, and observing the binding between the two, a transformant carrying the polynucleotide in an expressible manner and an antisense polynucleotide. The oligonucleotide is useful as a primer for synthesising the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a protein of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.
XX Sequence 158 AA;
SQ
Alignment Scores:
Pred. No.: 0.000172 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 7 Gaps: 0
US-10-664-025-43_COPY_53_103 (1-51) x ADB65520 (1-158)
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuTrpThrGlyAla 17
RESULT 11
ADC74338

ID ADC74338 standard; protein; 158 AA.
AC ADC74338;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human secreted protein - SEQ ID 971.
XX
XX antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
XX antidiabetic; immunosuppressive; dermatological; nephrotropic;
XX antiparkinsonian; neuroprotective; nontropic; antibacterial; virucide;
XX fungicide; antiparasitic; antiarteriosclerotic; vulnery; cytostatic;
XX haemopoietic; haematologic; anaemia; autoimmune disorder;
XX rheumatoid arthritis; inflammation; Grave's disease; diabetes;
XX systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
XX Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
XX cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
XX human.
XX
XX Homo sapiens.
XX
XX WO2003038063-A2.
XX
XX 08-MAY-2003.
XX
XX 19-MAR-2002; 2002WO-US008277.
XX
XX 21-MAR-2001; 2001US-0277340P.
XX 19-JUL-2001; 2001US-0306171P.
XX 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-430516/40.
XX
XX N-PSDB; ADC73723.
XX
XX New human secreted polypeptide for diagnosing, preventing or treating
XX hemopoietic or hematologic disorders (e.g. anemia), autoimmune
XX disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
XX atherosclerosis).
XX
XX Claim 16; SEQ ID NO 971; 2272pp; English.
XX
XX The invention relates to a novel human secreted polypeptide comprising a
XX defined sequence given in the specification. The polypeptide, nucleic
XX acid molecule, antibody, agonist or antagonist of the invention may be
XX useful for preparing a composition for diagnosing or treating a
XX haemopoietic or haematologic disorder such as anaemia, autoimmune
XX disorders such as rheumatoid arthritis, inflammation, Grave's disease,
XX diabetes, systemic lupus erythematosus or glomerulonephritis,
XX neurodegenerative disorders including Parkinson's disease and Alzheimer's
XX disease, wounds and hyperproliferative disorders including
XX atherosclerosis or cancer, as well as bacterial, viral, fungal or
XX parasitic infections. The polypeptide may also be used during gene
XX therapy procedures and for identifying a binding partner by contacting
XX the polypeptide with a binding partner and determining whether the
XX binding partner increases or decreases the activity of the polypeptide.
XX The current sequence is that of the human secreted protein of the
XX invention.

XX SQ Sequence 158 AA;

Alignment Scores:

Pred. No.:	0.000172	Length:	158
Score:	81.00	Matches:	17
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	88.0%	Indels:	0
DB:	7	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x ADC74338 (1-158)

QY 1 ATGCGGTAGAACTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT 51
DB |||||
1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17
RESULT 12
AAB38337
ID AAB38337 standard; protein; 159 AA.
XX
XX AAB38337;
XX
XX 31-JAN-2001 (first entry)
XX
XX Human secreted protein encoded by gene 17 clone HTLS08.
XX
XX Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
XX cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
XX nontropic; antibacterial; virucide; fungicide; ophthalmological; human;
XX vulnery; gene therapy; infection; secreted protein.
XX
XX Homo sapiens.
XX
XX WO200061623-A1.
XX
XX 19-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US008979.
XX
XX 09-APR-1999; 99US-0128693P.
XX 26-APR-1999; 99US-0130591P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;
XX Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;
XX Young PE;
XX
XX WPI; 2000-647418/62.
XX
XX New nucleic acid molecules encoding 62 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives.
XX
XX Claim 11; Page 608-609; 716pp; English.

XX Sequences AAB38321-B38396 represent the amino acid sequences of 62 human
XX secreted proteins encoded by the genes AAC69512-C69587. The genes and
XX proteins are useful for preventing, ameliorating or treating medical
XX conditions, e.g. by protein or gene therapy. The genes are isolated from
XX a range of human tissues disclosed in the specification. The nucleic
XX acids, proteins, antibodies and (ant)agonists are useful in the
XX diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
XX rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of
XX the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;
XX (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angioneurosis;
XX (f) nervous system disorders e.g. Alzheimer's disease; (g) infections
XX caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.
XX corneal infection. The polypeptides can also be used to aid wound healing
XX and epithelial cell proliferation, to prevent skin aging due to sunburn,
XX to maintain organs before transplantation, for supporting cell culture of
XX primary tissues, to regenerate tissues and in chemotaxis

XX SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	0.000173	Length:	159
Score:	81.00	Matches:	17
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	88.0%	Indels:	0
DB:	3	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x AAB38337 (1-159)

Oy 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGAATTGGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTyrThrGlyAla 17

RESULT 13

AAW14446
ID AAW14446 standard; protein; 250 AA.

XX AAW14446;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-MAY-1997 (first entry)
XX
DE CarB gene product.
XX
KW carbenapem; carR; carA; carB; carC; carD; carE; carF; carG; carH;
KW biosynthesis; antibiotic; beta-lactamase inhibitor; regulatory.

XX Pectobacterium carotovorum.

XX WO9532294-A1.

XX 30-NOV-1995.

XX 18-MAY-1995; 95WO-GB001125.

XX 20-MAY-1994; 94GB-00010142.

XX (UYWA-) UNIV WARWICK.

PA (UYNO-) UNIV NOTTINGHAM.

XX Salmond GPC, McGowan SJ, Sebahia M, Cox ARJ, Holden MTG;
PI Porter LE, Bycroft BW, Williams P, Stewart GSAB;

XX WPI; 1996-020587/02.

DR N-PSDB; AAT09700.

XX New genes encoding enzymes involved in carbenapem biosynthesis - useful
PT for isolating other carbenapem synthesising genes and producing
PT carbenapem in heterologous organisms.

XX Disclosure; Fig 4; 31pp; English.

XX AAW14445-52 are the protein products of the car genes (carA-H) which are
CC involved in the biosynthesis of carbenapem, an antibiotic which acts as a
CC beta-lactamase inhibitor. The car biosynthetic genes are regulated by the
CC carR gene product (AAW14444), a positive activator of the biosynthetic
CC genes which functions in trans. The car genes can be used to identify
CC homologous genes in bacteria and fungi, while products of the genes may
CC be used to raise antibodies used to detect clones (in a Streptomyces
CC library) that express cross-reactive protein. (Updated on 25-MAR-2003 to
CC correct PI field.) (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 250 AA;

Alignment Scores:			
Pred. No.:	76.4	Length:	250
Score:	47.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	80.0%	Mismatches:	0
Query Match:	51.1%	Indels:	0
DB:	2	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x AAW14446 (1-250)

Oy 9 AGAAGCTTGATGCTCCTCGCTGCTGTGAT 38

|||||

Db 162 GlnSerLeuAspAlaProArgCysValasp 171

RESULT 14

ABG79696

ABG79696 standard; protein; 624 AA.

ABG79696;

DT 15-NOV-2002 (first entry)

XX Tumour involved gene (TIG) splice variant protein, NV-27.

DE Human; splice variant; tumour-involved gene; TIG;

XX Pharmaceutical composition; cancer; diagnostic; tumour; gene therapy;
KW endothelial cell; cell differentiation; cell proliferation; apoptosis;
KW gene therapy.

XX Homo sapiens.

XX US2002086384-A1.

XX 04-JUL-2002.

XX 13-MAR-2001; 2001US-00805020.

PR 14-MAR-2000; 2000IL-00135402.

PR 16-MAY-2000; 2000IL-00136154.

XX (LEVI/) LEVINE Z.

PA (DAVI/) DAVID A.

PA (ROMA/) ROMANO C.

PA (BERN/) BERNSTEIN J.

XX Levine Z, David A, Romano C, Bernstein J;

WPI; 2002-635679/68.

DR N-PSDB; ABS65226.

XX Novel nucleic acid sequence, which is an alternative splicing variant of
PT tumor involved genes, useful for detecting cancer, predisposition to
PT cancer, for evaluating cancer state and in gene therapy for treating
PT cancer.

XX Claim 4; Page 94-95; 180pp; English.

XX The invention discloses isolated human nucleic acid alternative splicing
CC variants that are all tumour-involved genes (TIGs). The nucleic acids and
CC polypeptides are useful for determining the level of a nucleic acid or
CC polypeptide in a biological sample, for detecting a variant nucleic acid
CC or polypeptide sequence in a biological sample, for determining the level
CC of variant nucleic acid or polypeptide sequences in a biological sample
CC and for determining the ratio between the level of variant sequence in a
CC first biological sample and the level of the original sequence from which
CC the variant has been varied by alternative splicing in a second
CC biological sample and for raising antibodies. A pharmaceutical
CC composition comprising a carrier and the nucleic acid, is useful for
CC treating diseases (e.g. cancer) that can be ameliorated or cured by
CC increasing or decreasing the level of the encoded protein. The nucleic
CC acids are also useful for diagnostic purposes, especially for detecting
CC cancer or a predisposition to cancer, for evaluating the state or
CC aggressiveness of cancer disease, in basic research, for understanding
CC the physiological function of the original TIG, in targeting or
CC developing pharmaceuticals, for distinguishing various stages in the life
CC cycle of the same type of cells which may be helpful for the development
CC of pharmaceuticals for various cancer stages in which cell cycle is non-
CC normal, for determining mutations in tumour-involved genes and in gene
CC therapy. The polypeptides are useful for identifying compounds capable of
CC binding to the variant product and modulating its activity and for
CC modulating endothelial differentiation and proliferation, as well as to
CC modulate apoptosis either ex vivo or in vivo. The sequences presented in
CC ABG796700-ABG79705 are the new variants (NV) 1-36 proteins of the TIGs
CC disclosed

XX Sequence 624 AA;

Alignment Scores:

Pred. No.:

79.3

Length:

624

Score: 47.00 Matches: 9
Percent Similarity: 76.9% Conservative: 1
Best Local Similarity: 69.2% Mismatches: 3
Query Match: 51.1% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x ABG79696 (1-624)

QY 51 AGCACCGGTCCAATCAAGACAGCGAGCATCAAGAC 13
||||| ||| ||||||||| |||||:::
Db 518 SerThrAspGluAnThrAspSerGluGluHisGlnGlu 530

RESULT 15

ADD14149

ID ADD14149 standard; protein; 633 AA.

AC ADD14149;

DT 01-JAN-2004 (first entry)

DE Human src biomarker polypeptide SEQ ID NO:338.

XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human.

OS Homo sapiens.

FN WO2003062395-A2.

XX 31-JUL-2003.

XX 17-JAN-2003; 2003WO-US001981.

XX 18-JAN-2002; 2002US-0350061P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Huang F, Fairchild CR, Lee FY, Shaw P;

DR WPI; 2003-636735/60.

DR N-PSDB; ADD14751.

XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.

XX Claim 10; SEQ ID NO 338; 139pp; English.

XX The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized

CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.

XX

SQ Sequence 633 AA;

Alignment Scores:

Pred. No.: 79.4 Length: 633
Score: 47.00 Matches: 9
Percent Similarity: 76.9% Conservative: 1
Best Local Similarity: 69.2% Mismatches: 3
Query Match: 51.1% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x ADD14149 (1-633)

QY 51 AGCACCGGTCCAATCAAGACAGCGAGCATCAAGAC 13
||||| ||| ||||||||| |||||:::

Db 518 SerThrAspGluAnThrAspSerGluGluHisGlnGlu 530

Search completed: May 9, 2006, 10:04:46
Job time : 13.1 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 9, 2006, 10:08:38 ; Search time 0.78 Seconds
(without alignments)
1621.715 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103
Perfect score: 92
Sequence: 1 atggcgtagaagttctgat.....tcttgattggaccgggtgct 51

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlip
-Q=/abss/ABSSWEB_spool/US10664025/runat_08052006_173459_28431/app_query.fasta_1
-DB=Issued_Patents_AA -QFMT=faстан -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=ext -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02p -USER=US10664025_@CGN1_1_101_@runat_08052006_173459_28431
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORE=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	81	88.0	US-09-621-976-5223
2	81	88.0	Sequence 5223, Ap
3	81	88.0	Sequence 3903, Ap
4	81	88.0	Sequence 5224, Ap
5	47	51.1	US-10-104-047-3674
6	47	51.1	Sequence 3674, Ap
7	46	50.0	US-09-949-016-10511
8	46	50.0	Sequence 10511, A
9	45	48.9	US-09-252-991A-23223
10	45	48.9	Sequence 23223, A
11	45	48.9	US-08-132-649-2
12	45	48.9	Sequence 2, Appli
13	45	48.9	Sequence 4, Appli
14	45	48.9	Sequence 2, Appli
15	45	48.9	Sequence 20495, A

US-10-664-025-43_COPY_53_103 (1-51) x US-09-621-976-5223 (1-106)

QY 1 ATGGCGTAGAAGTTCTGATCTCTCGCTGCTTGAATTCGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla 17

13	44	47.8	33	2	US-10-290-579A-165	Sequence 165, App
14	44	47.8	86	2	US-10-290-579A-248	Sequence 248, App
15	44	47.8	180	2	US-09-328-352-5855	Sequence 5855, Ap
16	44	47.8	223	2	US-09-270-767-32256	Sequence 32256, A
17	44	47.8	347	1	US-09-004-502-1	Sequence 1, Appli
18	44	47.8	347	2	US-09-360-125-1	Sequence 1, Appli
19	44	47.8	810	1	US-08-820-170A-34	Sequence 34, Appl
20	44	47.8	810	2	US-09-055-899-34	Sequence 34, Appl
21	44	47.8	810	2	US-09-273-565-34	Sequence 34, Appl
22	44	47.8	810	2	US-09-565-538-34	Sequence 34, Appl
23	44	47.8	810	2	US-09-661-468-34	Sequence 34, Appl
24	44	47.8	810	2	US-09-976-165-34	Sequence 34, Appl
25	43	46.7	33	2	US-10-290-579A-166	Sequence 166, App
26	43	46.7	86	2	US-10-290-579A-248	Sequence 248, App
27	43	46.7	547	2	US-09-489-039A-13843	Sequence 13843, A
28	43	46.7	743	2	US-09-248-796A-15523	Sequence 15523, A
29	42	45.7	27	2	US-10-290-579A-171	Sequence 171, App
30	42	45.7	28	2	US-10-290-579A-168	Sequence 168, App
31	42	45.7	80	2	US-10-290-579A-240	Sequence 240, App
32	42	45.7	201	2	US-09-902-540-11728	Sequence 11728, A
33	42	45.7	237	2	US-09-252-991A-17927	Sequence 17927, A
34	42	45.7	425	2	US-09-252-991A-26326	Sequence 26326, A
35	42	45.7	446	2	US-09-489-039A-14283	Sequence 14283, A
36	42	45.7	457	2	US-09-252-991A-17452	Sequence 17452, A
37	42	45.7	1210	2	US-10-771-708-10	Sequence 10, Appl
38	42	45.7	1833	2	US-08-479-722B-2	Sequence 2, Appli
39	42	45.7	1833	2	US-09-592-685-2	Sequence 2, Appli
40	42	45.7	1833	4	PCT-US95-02251-18	Sequence 18, Appl
41	42	45.7	3571	2	US-09-911-842A-2	Sequence 2, Appli
42	42	45.7	3594	2	US-09-911-842A-4	Sequence 4, Appli
43	41.5	45.1	439	2	US-09-409-096-6	Sequence 6, Appli
44	41.5	45.1	443	1	US-08-833-963C-2	Sequence 2, Appli
45	41.5	45.1	443	2	US-08-980-514-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-621-976-5223
; Sequence 5223, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5223
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-621-976-5223

Alignment Scores:			
Pred. No.:	Score:	Length:	Matches:
5.78e-06	81.00	106	17
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	88.0%	Indels:	0
DB:	2	Gaps:	0
US-10-664-025-43_COPY_53_103 (1-51) x US-09-621-976-5223 (1-106)			
QY	1	ATGGCGTAGAAGTTCTGATCTCTCGCTGCTTGAATTCGACCGGTGCT	51
Db	1	MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAla	17

RESULT 2
US-09-621-976-3903
; Sequence 3903, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3903
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-621-976-3903

Alignment Scores:
Pred. No.: Length: 153
Score: 6.06e-06 Matches: 17
Percent Similarity: 81.00% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 88.0% Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-621-976-3903 (1-153)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAla 17

RESULT 3
US-09-621-976-5222
; Sequence 5222, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5222
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-621-976-5222

Alignment Scores:
Pred. No.: Length: 158
Score: 6.09e-06 Matches: 17
Percent Similarity: 81.00% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 88.0% Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-621-976-5222 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||

Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAla 17

RESULT 4
US-10-104-047-3674
; Sequence 3674, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674

Alignment Scores:
Pred. No.: Length: 158
Score: 6.09e-06 Matches: 17
Percent Similarity: 81.00% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 88.0% Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-104-047-3674 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAla 17

RESULT 5
US-08-737-825-4
; Sequence 4, Application US/08737825
; Patent No. 5871922
; GENERAL INFORMATION:
; APPLICANT: SALMOND, GEORGE PEACOCK COPELAND
; APPLICANT: MCGOWAN, SIMON JAMES
; APPLICANT: SEBAIHIA, MOHAMMED
; APPLICANT: COX, ANTHONY RICHARD JOHN
; APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY
; APPLICANT: PORTER, LAUREN ELIZABETH
; APPLICANT: BYCROFT, BARRIE WALSHAW
; APPLICANT: STEWART, GORDON SIDNEY ANDERSON BIRNIE
; TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,825
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 1009-0105P
; TELECOMMUNICATION INFORMATION:


```
;
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Erwinia carotovora
;
US-08-737-825-4
Alignment Scores:
Pred. No.: 6.82 Length: 250
Score: 17.00 Matches: 8
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 80.0% Mismatches: 0
Query Match: 51.1% Indels: 0
DB: 1 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-08-737-825-4 (1-250)
QY 9 AGAAGTCTTGATGCTCTCGCTGCTTGAT 38
Db 162 GlnSerLeuAspAlaProArgCysValasp 171

RESULT 6
US-09-949-016-10511
; Sequence 10511, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10511
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Human
;
US-09-949-016-10511
Alignment Scores:
Pred. No.: 7.75 Length: 668
Score: 47.00 Matches: 9
Percent Similarity: 76.9% Conservative: 1
Best Local Similarity: 69.2% Mismatches: 3
Query Match: 51.1% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-949-016-10511 (1-668)
QY 51 AGCACGGTCCAAATCAAGACAGCGAGGAGCATCAAGAC 13
Db 553 SerThrAspGluAsnThrAspSerGluGluHisGlnGlu 565

RESULT 7
US-09-252-991A-23223
; Sequence 23223, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23223
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-23223
Alignment Scores:
Pred. No.: 10 Length: 209
Score: 46.00 Matches: 8
Percent Similarity: 92.3% Conservative: 4
Best Local Similarity: 61.5% Mismatches: 1
Query Match: 50.0% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-252-991A-23223 (1-209)
QY 49 CACCGTCCAAATCAAGACAGCGAGGAGCATCAAGACTT 11
Db 150 HisArgSerArgAlaArgArgGlyArgThrArgLeu 162

RESULT 8
US-09-252-991A-23254
; Sequence 23254, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23254
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-23254
Alignment Scores:
Pred. No.: 10.7 Length: 342
Score: 46.00 Matches: 8
Percent Similarity: 92.3% Conservative: 4
Best Local Similarity: 61.5% Mismatches: 1
Query Match: 50.0% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-252-991A-23254 (1-342)
QY 49 CACCGTCCAAATCAAGACAGCGAGGAGCATCAAGACTT 11
Db 150 HisArgSerArgAlaArgArgGlyArgThrArgLeu 162

RESULT 9
US-08-132-649-2
; Sequence 2, Application US/08132649
; Patent No. 5585462
; GENERAL INFORMATION:
; APPLICANT: London, Constantine
```

APPLICANT: Greenberg, Andrew S.
APPLICANT: Kimmel, Alan R.
APPLICANT: Egan, John J.
TITLE OF INVENTION: CLONING OF PERILIPIN PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSES: Townsend and Townsend Hourie and Crew
STREET: Stuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/132,649
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-145-1
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-132-649-2

Alignment Scores:
Pred. No.: 16.9 Length: 517
Score: 45.00 Matches: 9
Percent Similarity: 69.2% Conservative: 0
Best Local Similarity: 69.2% Mismatches: 4
Query Match: 48.9% Indels: 0
DB: 1 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-08-132-649-2 (1-517)

QY 6 GTTAGAGCTCTGATGCTCTCGCTGCTTGTGATTGGAC 44
||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 456 ValArgGlyLeuSerAlaProSerCysProAspLeuAsp 468

RESULT 10
US-08-764-343-4
Sequence 4, Application US/08764343
Patent No. 5739009
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL ADIPOCYTE-SPECIFIC
DIFFERENTIATION-RELATED PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSES: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/764,343
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0167 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1172433
US-08-764-343-4

Alignment Scores:
Pred. No.: 16.9 Length: 517
Score: 45.00 Matches: 9
Percent Similarity: 69.2% Conservative: 0
Best Local Similarity: 69.2% Mismatches: 4
Query Match: 48.9% Indels: 0
DB: 1 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-08-764-343-4 (1-517)

QY 6 GTTAGAGCTCTGATGCTCTCGCTGCTTGTGATTGGAC 44
||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 456 ValArgGlyLeuSerAlaProSerCysProAspLeuAsp 468

RESULT 11
US-08-767-579-2
Sequence 2, Application US/08767579
Patent No. 6074842
GENERAL INFORMATION:
APPLICANT: Londres, Constantine
APPLICANT: Greenberg, Andrew S.
APPLICANT: Kimmel, Alan R.
APPLICANT: Egan, John J.
TITLE OF INVENTION: CLONING OF PERILIPIN PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSES: Townsend and Townsend Hourie and Crew
STREET: Stuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,579
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-145-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:

```

; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 165
; LENGTH: 33
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: isolate=94IN476.104; gene=vpu
US-10-290-579A-165

Alignment Scores:
Pred. No.:          17.8      Length:         33
Score:              44.00     Matches:        7
Percent Similarity: 85.7%    Conservative:   5
Best Local Similarity: 50.0% Mismatches:      2
Query Match:       47.8%    Indels:         0
DB:                2        Gaps:           0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-290-579A-165 (1-33)

Qy      4 GCCTTAGAGTCTTGATGCTCCTCGCTGTCTGATTGGACC 45
||||| ||| :::::|||||::::::::::|||
Db      16 AlAlAuLeUleValAlaLeuAlaIleValTrpThr 29

RESULT 14
US-10-290-579A-248
; Sequence 248, Application US/10290579A
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B I
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 248
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94IN476.104; gene=vpu
US-10-290-579A-248

Alignment Scores:
Pred. No.:          20.2      Length:         86
Score:              44.00     Matches:        7
Percent Similarity: 85.7%    Conservative:   5
Best Local Similarity: 50.0% Mismatches:      2
Query Match:       47.8%    Indels:         0
DB:                2        Gaps:           0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-290-579A-248 (1-86)

Qy      4 GCCTTAGAGTCTTGATGCTCCTCGCTGTCTGATTGGACC 45
||||| ||| :::::|||||::::::::::|||
Db      16 AlAlAuLeUleValAlaLeuAlaIleValTrpThr 29

RESULT 15
US-09-328-352-5855
; Sequence 5855, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RE
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTI
; FILE REFERENCE: GTC99-03PAA

```

```
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5855
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5855

Alignment Scores:
Pred. No.:      22.2      Length:      180
Score:          44.00     Matches:      7
Percent Similarity: 100.0% Conservative: 6
Best Local Similarity: 53.8% Mismatches: 0
Query Match:     47.8% Indels:      0
DB:              2      Gaps:       0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-328-352-5855 (1-180)

QY      10 GAAGTCTTGATGCTCTCGCTGCTTGTGATTGGACGGT 48
Db      140 LysValLeuIleuIleSerLeuIleTrpSerGly 152

Search completed: May 9, 2006, 10:10:06
Job time : 5.9 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:09:27 ; Search time 1.81 Seconds
(without alignments)
3531.927 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103
Perfect score: 92
Sequence: 1 atggcgtagaagcttgat.....tcttgattggaccggtgct 51

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1867569 segs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-Q=/abss/ABSSWEB.spool/US10664025/runat_08052006_173502_28483/app_query.fasta_1
-DB=Published Applications AA Main -OFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss07
-USER=US10664025 @CGN 1.1.198 @runat_08052006_173502_28483 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	81	88.0	106	3	US-09-731-872-265	Sequence 265, App
2	81	88.0	106	3	US-09-876-997-265	Sequence 265, App
3	81	88.0	106	5	US-10-643-836-265	Sequence 265, App
4	81	88.0	158	3	US-09-731-872-260	Sequence 260, App
5	81	88.0	158	3	US-09-876-997-260	Sequence 260, App
6	81	88.0	158	4	US-10-050-704-113	Sequence 113, App
7	81	88.0	158	4	US-10-104-047-3674	Sequence 3674, App
8	81	88.0	158	4	US-10-467-535-4	Sequence 4, Appli
9	81	88.0	158	4	US-10-798-512-113	Sequence 113, App
10	81	88.0	158	5	US-10-467-046-3	Sequence 3, Appli
11	81	88.0	158	5	US-10-643-836-260	Sequence 260, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Alignment Scores:	0.000226	Length:	106
Pred. No.:	81.00	Matches:	17
Score:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	88.0%	Indels:	0
DB:	3	Gaps:	0

ALIGNMENTS

RESULT 1
US-09-731-872-265
; Sequence 265, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueloret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-731-872-265

12	81	88.0	158	5	US-10-472-533-464	Sequence 464, App
13	50	54.3	75	4	US-10-425-115-251583	Sequence 251583,
C 14	47	51.1	110	4	US-10-425-115-206002	Sequence 206002,
C 15	47	51.1	624	3	US-09-805-020-63	Sequence 63, Appl
C 16	47	51.1	633	4	US-10-087-192-1800	Sequence 1800, Ap
C 17	47	51.1	633	4	US-10-116-275-175	Sequence 175, App
18	47	51.1	704	4	US-10-424-599-161519	Sequence 161519,
19	47	51.1	709	4	US-10-424-599-161541	Sequence 161541,
20	47	51.1	714	4	US-10-425-114-48396	Sequence 48396, A
21	47	51.1	996	4	US-10-282-122A-50437	Sequence 50437, A
22	46	50.0	102	4	US-10-425-114-71786	Sequence 71786, A
23	46	50.0	138	4	US-10-425-114-66911	Sequence 66911, A
24	46	50.0	345	4	US-10-369-493-8151	Sequence 8151, Ap
25	46	50.0	694	4	US-10-156-761-11175	Sequence 11775, A
26	46	50.0	709	5	US-10-739-930-5790	Sequence 5790, Ap
27	46	50.0	862	4	US-10-281-478-3	Sequence 3, Appli
28	46	50.0	862	4	US-10-464-368-90	Sequence 90, Appl
29	46	50.0	862	4	US-10-464-368-91	Sequence 91, Appl
30	46	50.0	864	4	US-10-464-368-92	Sequence 92, Appl
31	45	48.9	104	3	US-09-833-245-1859	Sequence 1859, Ap
32	45	48.9	114	3	US-09-833-245-1860	Sequence 1860, Ap
33	45	48.9	114	4	US-10-425-115-259901	Sequence 259901,
34	45	48.9	179	4	US-10-017-161-974	Sequence 974, App
C 35	45	48.9	209	4	US-10-437-963-153479	Sequence 153479,
36	45	48.9	232	4	US-10-425-115-259897	Sequence 259897,
C 37	45	48.9	354	5	US-10-732-923-7962	Sequence 7962, Ap
38	45	48.9	406	4	US-10-028-072-82	Sequence 82, Appl
39	45	48.9	406	4	US-10-140-808-82	Sequence 82, Appl
40	45	48.9	406	4	US-10-121-049-82	Sequence 82, Appl
41	45	48.9	406	4	US-10-123-904-82	Sequence 82, Appl
42	45	48.9	406	4	US-10-140-470-82	Sequence 82, Appl
43	45	48.9	406	4	US-10-175-746-82	Sequence 82, Appl
44	45	48.9	406	4	US-10-176-918-82	Sequence 82, Appl
45	45	48.9	406	4	US-10-176-921-82	Sequence 82, Appl

US-10-664-025-43_COPY_53_103 (1-51) x US-09-731-872-265 (1-106)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 2

US-09-876-997-265
; Sequence 265, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-876-997-265

Alignment Scores:
Pred. No.: 0.000226 Length: 106
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
Db: 3 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-876-997-265 (1-106)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 3

US-10-643-836-265
; Sequence 265, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-10-643-836-265

Alignment Scores:

Pred. No.: 0.000226 Length: 106
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
Db: 5 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-643-836-265 (1-106)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 4

US-09-731-872-260
; Sequence 260, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-731-872-260

Alignment Scores:

Pred. No.: 0.000231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
Db: 3 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-731-872-260 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 5

US-09-876-997-260
; Sequence 260, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872

```
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-876-997-260

Alignment Scores:
Pred. No.: 0.000231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-876-997-260 (1-158)
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 6
US-10-050-704-113
; Sequence 113, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-113

Alignment Scores:
Pred. No.: 0.000231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-050-704-113 (1-158)
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 7
US-10-104-047-3674
; Sequence 3674, Application US/10104047
```

```
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674

Alignment Scores:
Pred. No.: 0.000231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-104-047-3674 (1-158)
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 8
US-10-467-535-4
; Sequence 4, Application US/10467535
; Publication No. US20040146970A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; YAO, Monique G.
; APPLICANT: ISON, Craig H.; LU, Yan
; APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
; APPLICANT: BAUGHN, Mariah R.; DING, Li
; APPLICANT: XU, Yuming; GIETZEN, Kimberly J.
; APPLICANT: TANG, Tom Y.; LAL, Preeti G.
; APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
; APPLICANT: LU, Dying Aina M.; RICHARDSON, Thomas W.
; APPLICANT: TRAN, Uyen K.; KHARE, Reena
; APPLICANT: CHAWLA, Narinder K.
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-0903 USN
; CURRENT APPLICATION NUMBER: US/10/467,535
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/US02/03715
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/268,111
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/271,175
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,503
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/274,552
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5284076CD1
US-10-467-535-4

Alignment Scores:
Pred. No.: 0.000231 Length: 158
```

```
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-467-535-4 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrrThrGlyAla 17

RESULT 9
US-10-798-512-113
; Sequence 113, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT FILING DATE: 2004-03-12
; PRIOR FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-798-512-113

Alignment Scores:
Pred. No.: 0.00231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-798-512-113 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrrThrGlyAla 17

RESULT 10
US-10-467-046-3
; Sequence 3, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSPf3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 158
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(17)
US-10-467-046-3

Alignment Scores:
Pred. No.: 0.000231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-467-046-3 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrrThrGlyAla 17

RESULT 11
US-10-643-836-260
; Sequence 260, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-10-643-836-260

Alignment Scores:
Pred. No.: 0.000231 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-643-836-260 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTGGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrrThrGlyAla 17

RESULT 12
US-10-472-533-464
; Sequence 464, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS906PCT
; CURRENT APPLICATION NUMBER: US/10/472,533
```



```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 206002
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_119456C.1.pep
US-10-425-115-206002

Alignment Scores:
Pred. No.:      83.7      Length:      110
Score:          47.00     Matches:      10
Percent Similarity: 80.0%   Conservative:  2
Best Local Similarity: 66.7% Mismatches:    3
Query Match:      51.1%   Indels:       0
DB:               4       Gaps:         0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-425-115-206002 (1-110)

Qy  47  CCGTCCAAATCAAGACGAGGAGCATCAAGACTTCTTAACGCC 3
    |||::|||::|||::|||::|||::|||::|||::|||
Db   26  Prol EuglinileysSerAlaArgproIleLysThrValGlyAla 40

RESULT 15
US-09-805-020-63
; Sequence 63, Application US/09805020
; Publication No. US20020086384A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
; FILE REFERENCE: 2786-0168P
; CURRENT APPLICATION NUMBER: US/09/805,020
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-020-63

Alignment Scores:
Pred. No.:      92.3      Length:      624
Score:          47.00     Matches:      9
Percent Similarity: 76.9%   Conservative:  1
Best Local Similarity: 69.2% Mismatches:    3
Query Match:      51.1%   Indels:       0
DB:               3       Gaps:         0

US-10-664-025-43_COPY_53_103 (1-51) x US-09-805-020-63 (1-624)

Qy   51  AGCACCGGTCCCAATCAAGACGAGGAGCATCAAGAC 13
    |||::|||::|||::|||::|||::|||::|||::|||
Db   518  SerThrAspGluAsnThrAspSerGluGluHisGlnGlu 530

Search completed: May 9, 2006, 10:13:13
Job time : 10.05 secs

```

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:10:23 ; Search time 0.53 Seconds
(without alignments)
1336.144 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 92

Sequence: 1 atggcgctagaagtcttgat.....tcttgattggaccggtgct 51

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 470810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp

-Q=/abss/ABSSWEB_spool/US10664025/runat_08052006_173505_28538/app_query.fasta.1

-DB=Published Applications AA New -QPMF=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62

-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0

-MAXLEN=200000000 -HOST=abss02p

-USER=US10664025 @CGN 1 1 47 @runat_08052006_173505_28538 -NCPU=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:

1: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pep1.*

2: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

3: /SIDSS/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

4: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

5: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep.*

6: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep.*

7: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep.*

8: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep.*

9: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep.*

10: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep.*

11: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep.*

12: /SIDSS/ptodata/1/pubpaa/US16_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	88.0	143	9	US-10-475-075-787 Sequence 787, App
2	81	88.0	148	9	US-10-475-075-788 Sequence 788, App
3	81	88.0	158	11	US-11-072-512-3674 Sequence 3674, App
C 4	47	51.1	633	9	US-10-501-035-338 Sequence 338, App
C 5	47	51.1	703	9	US-10-784-004-1234 Sequence 1234, App

6	45	48.9	104	11	US-11-264-096-1859	Sequence 1859, App
7	45	48.9	104	11	US-11-264-096-1860	Sequence 1860, App
C 8	45	48.9	116	11	US-11-096-568A-8673	Sequence 8673, App
C 9	45	48.9	126	11	US-11-096-568A-8672	Sequence 8672, App
C 10	45	48.9	131	11	US-11-096-568A-8671	Sequence 8671, App
C 11	45	48.9	163	11	US-11-098-686-10809	Sequence 10809, App
12	45	48.9	406	9	US-10-131-826A-82	Sequence 82, Appl
13	45	48.9	406	9	US-10-973-115B-82	Sequence 82, Appl
14	45	48.9	406	9	US-10-137-873A-82	Sequence 82, Appl
15	45	48.9	406	9	US-10-152-370-82	Sequence 82, Appl
16	45	48.9	406	11	US-11-290-153-82	Sequence 82, Appl
17	44	47.8	347	9	US-10-821-234-1379	Sequence 1379, App
18	44	47.8	810	11	US-11-192-813-2	Sequence 2, Appl
19	44	47.8	810	11	US-11-192-813-4	Sequence 4, Appl
20	44	47.8	810	11	US-11-192-813-6	Sequence 6, Appl
C 21	43	46.7	266	11	US-11-096-568A-17804	Sequence 17804, A
C 22	43	46.7	319	11	US-11-096-568A-17803	Sequence 17803, A
C 23	43	46.7	422	9	US-10-942-698-3	Sequence 3, Appl
C 24	43	46.7	422	11	US-11-022-562-234	Sequence 234, App
C 25	43	46.7	428	11	US-11-004-399-924	Sequence 924, App
C 26	42.5	46.2	399	11	US-11-087-295-169	Sequence 169, App
C 27	42	45.7	92	9	US-10-667-295-169	Sequence 169, App
C 28	42	45.7	92	11	US-11-172-740-1918	Sequence 1918, App
C 29	42	45.7	92	11	US-11-172-740-2404	Sequence 2404, App
C 30	42	45.7	436	11	US-11-116-939-9	Sequence 9, Appl
C 31	42	45.7	455	11	US-11-096-568A-27476	Sequence 27476, A
C 32	42	45.7	469	11	US-11-096-568A-27475	Sequence 27475, A
C 33	42	45.7	489	11	US-11-079-463-6945	Sequence 6945, App
C 34	42	45.7	522	11	US-11-096-568A-27474	Sequence 27474, A
C 35	42	45.7	824	11	US-11-116-939-11	Sequence 11, Appl
C 36	42	45.7	1210	11	US-11-191-374-10	Sequence 10, Appl
C 37	42	45.7	1210	11	US-11-191-375-10	Sequence 10, Appl
C 38	42	45.7	1210	11	US-11-191-588-10	Sequence 10, Appl
C 39	42	45.7	1272	9	US-10-501-035-313	Sequence 313, App
C 40	42	45.7	2312	11	US-11-126-313-34	Sequence 34, App
C 41	42	45.7	3568	9	US-10-453-372-194	Sequence 194, App
C 42	42	45.7	3570	9	US-10-453-372-178	Sequence 178, App
C 43	42	45.7	3570	9	US-10-453-372-196	Sequence 196, App
C 44	42	45.7	3570	9	US-10-453-372-198	Sequence 198, App
C 45	42	45.7	3570	9	US-10-453-372-200	Sequence 200, App

ALIGNMENTS

RESULT 1

US-10-475-075-787
; Sequence 787, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 787
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-475-075-787

Alignment Scores:

Pred. No.: 3.69e-06 Length: 143

Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 9 Gaps: 0
US-10-664-025-43_COPY_53_103 (1-51) x US-10-475-075-787 (1-143)
QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 2
US-10-475-075-788
; Sequence 788, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 788
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-475-075-788

Alignment Scores:
Pred. No.: 3,71e-06 Length: 148
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-475-075-788 (1-148)
QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 3
US-11-072-512-3674
; Sequence 3674, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3674
Alignment Scores:
Pred. No.: 3,73e-06 Length: 158
Score: 81.00 Matches: 17
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 88.0% Indels: 0
DB: 11 Gaps: 0
US-10-664-025-43_COPY_53_103 (1-51) x US-11-072-512-3674 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTGACCGGTGCT 51
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 4
US-10-501-035-338
; Sequence 338, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 338
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-338

Alignment Scores:
Pred. No.: 4,23 Length: 633
Score: 47.00 Matches: 9
Percent Similarity: 76.9% Conservative: 1
Best Local Similarity: 69.2% Mismatches: 3
Query Match: 51.1% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-501-035-338 (1-633)
QY 51 AGCACCGTCCAAATCAAGACAGCAGGAGCATCAAGAC 13
Db 518 SerThrAspGluAsnThrAspSerGluGluHisGlnGlu 530

RESULT 5
US-10-784-004-1234
; Sequence 1234, Application US/10784004
; Publication No. US20060084066A1


```

; Sequence 8672, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR FILING DATE: 2003-10-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8672
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(126)
; OTHER INFORMATION: Ceres Seq. ID no. 13589014
US-11-096-568A-8672

Alignment Scores:
Pred. No.:      8 07      Length:      126
Score:          45.00     Matches:       7
Percent Similarity: 75.0%     Conservative: 2
Best Local Similarity: 58.3%   Mismatches: 3
Query Match:     48.9%     Indels:    0
DB:              11       Gaps:      0

US-10-664-025-43_COPY_53_103 (1-51) x US-11-096-568A-8672 (1-126)

QY 45 GGTCCAAATCAAGACGACGAGGAGCATCAAGACTTC 10
      ||||| ||||| ||||| ||||| ||||| |||||
Db 115 GlyProAsnGlnGluAspGluSerGluGlnAspTyr 126

RESULT 10
US-11-096-568A-8671
; Sequence 8671, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8671
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(131)
; OTHER INFORMATION: Ceres Seq. ID no. 13589013
US-11-096-568A-8671

Alignment Scores:
Pred. No.:      8 1      Length:      131
Score:          45.00     Matches:       7
Percent Similarity: 75.0%     Conservative: 2
Best Local Similarity: 58.3%   Mismatches: 3
Query Match:     48.9%     Indels:    0
DB:              11       Gaps:      0

US-10-664-025-43_COPY_53_103 (1-51) x US-11-096-568A-8671 (1-131)

QY 45 GGTCCAAATCAAGACGACGAGGAGCATCAAGACTTC 10
      ||||| ||||| ||||| ||||| ||||| |||||
Db 120 GlyProAsnGlnGluAspGluSerGluGlnAspTyr 131

RESULT 11
US-11-098-686-10809
; Sequence 10809, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10809
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10809

Alignment Scores:
Pred. No.:      8 28      Length:      163
Score:          45.00     Matches:       6
Percent Similarity: 85.7%     Conservative: 6
Best Local Similarity: 42.9%   Mismatches: 2
Query Match:     48.9%     Indels:    0
DB:              11       Gaps:      0

US-10-664-025-43_COPY_53_103 (1-51) x US-11-098-686-10809 (1-163)

QY 7 TTAGAAGTCTTGATGCTCCTCGCTCTTGATTGGACCGGT 48
      |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 144 LeuGlnAlaLeuLeuIleSerIleLeuTyrTrpThrGly 157

RESULT 12
US-10-131-826A-82
; Sequence 82, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17

```

```
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 82
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-82

Alignment Scores:
Pred. No.: 9.1 Length: 406
Score: 45.00 Matches: 7
Percent Similarity: 64.3% Conservative: 2
Best Local Similarity: 50.0% Mismatches: 5
Query Match: 48.9% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-131-826A-82 (1-406)
QY 9 AGAAGTCTTGATGCTCTCGTCTGTTGATTGGACCGGTGC 50
Db 125 ArgLysArgGluLysProTyrCysLeuAspIleAspGluCys 138

RESULT 13
US-10-973-115B-82
; Sequence 82, Application US/10973115B
; Publication No. US20060040351A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: 39870-3330R1C300C1
; CURRENT APPLICATION NUMBER: US/10/973,115B
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 10/145,747
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/028,072
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/581,742
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: PCT/US00/05746
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/135,736
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 60/123,090
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 82
```

```
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-973-115B-82

Alignment Scores:
Pred. No.: 9.1 Length: 406
Score: 45.00 Matches: 7
Percent Similarity: 64.3% Conservative: 2
Best Local Similarity: 50.0% Mismatches: 5
Query Match: 48.9% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-973-115B-82 (1-406)
QY 9 AGAAGTCTTGATGCTCTCGTCTGTTGATTGGACCGGTGC 50
Db 125 ArgLysArgGluLysProTyrCysLeuAspIleAspGluCys 138

RESULT 14
US-10-137-873A-82
; Sequence 82, Application US/10137873A
; Publication No. US20060084138A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C149
; CURRENT APPLICATION NUMBER: US/10/137,873A
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 82
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-873A-82
```

```
Alignment Scores:
Pred. No.:      9.1      Length:      406
Score:          45.00    Matches:      7
Percent Similarity: 64.3%  Conservative: 2
Best Local Similarity: 50.0%  Mismatches: 5
Query Match:    48.9%    Indels:      0
DB:             9      Gaps:        0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-137-873A-82 (1-406)

QY      9  AGAAGTCTTGATGCTCGCTGCTGTTGATTGGACCGGTGC 50
      |||  ::  |||  |||||:::|  |||
Db      125  ArgLysArgGluLysProTyrCysLeuAspIleAspGluCys 138

RESULT 15
US-10-152-370-82
; Sequence 82, Application US/10152370
; Publication NO. US20060084139A1
; GENERAL INFORMATION:
; APPLICANT: Bersini, Kevin P.
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC407
; CURRENT APPLICATION NUMBER: US/10/152,370
; PRIOR APPLICATION DATE: 2002-05-21
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 82
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-152-370-82

Alignment Scores:
Pred. No.:      9.1      Length:      406
Score:          45.00    Matches:      7
Percent Similarity: 64.3%  Conservative: 2
Best Local Similarity: 50.0%  Mismatches: 5
Query Match:    48.9%    Indels:      0
DB:             9      Gaps:        0

US-10-664-025-43_COPY_53_103 (1-51) x US-10-152-370-82 (1-406)

QY      9  AGAAGTCTTGATGCTCGCTGCTGTTGATTGGACCGGTGC 50
      |||  ::  |||  |||||:::|  |||
Db      125  ArgLysArgGluLysProTyrCysLeuAspIleAspGluCys 138
```

Search completed: May 9, 2006, 10:11:25
Job time : 3.65 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 9, 2006, 10:04:56 ; Search time 0.48 Seconds
(without alignments)
3066.909 Million cell updates/sec

Title: US-10-664-025-43 COPY 53 103

Perfect score: 92

Sequence: 1 atggcgtagaagtcttgat.....tcttgatttggaccggttact 51

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5	Ygapop 10.0 , Ygapext 0.5
---------------------------	---------------------------

Egapop 6.0 , Egapext 7.0

Delop	6.0	Delext	7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL=framed_n2p.model -DEV=xlh
-Q=/abs/ABSSWE_spool/US10664025/runat_08052006.173455_28373/app_query.fasta_1
-DB=P -R -QMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bic -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DUALIGN=200 -THR1 SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=5 -MODE=LOCAL
-OUTFWT=pfo -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs07
-USER=US10664025 -CGN=1.35 -@runat_08052006.173455_28373 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGPOP=6 -FCGPEXT=7
YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : PIR 80: *

1: _pir1:*

```

1: pir1:
2: pir2: *

```

3: pir3: *

4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query			DB	ID	Description
			Match	Length	%			
C	1	50	54.3	334	2	H69148	hypothetical prote	
	2	47	51.1	274	2	S17804	hypothetical prote	
	3	47	51.1	633	2	A36353	DNA repair protein	
	4	46	50.0	710	2	T13458	hypothetical prote	
	5	46	50.0	862	1	QRMSLD	LDL receptor precu	
	6	45	48.9	101	2	S60434	probable membrane	
	7	45	48.9	309	2	T26440	hypothetical prote	
	8	45	48.9	517	2	A94913	perilipin A - rat	
	9	45	48.9	570	2	A83118	probable ATP-bindin	
	10	45	48.9	553	2	D69815	conserved hypothet	
C	11	44	47.8	347	2	T08826	secretory carrier	
	12	44	47.8	690	2	T47478	hypothetical prote	
	13	44	47.8	810	2	T10756	Nel-homolog protei	
	14	43	46.7	55	1	FECLCE	ferredoxin 2[4Fe-4	

ALIGNMENTS

RESULT 1

H69148
Hypothetical protein MTH378 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: H69148
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge,
Q.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivan,
K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H:
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: H69148
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

```

A; gene: wtn370
C; Superfamily: conserved hypothetical protein MTH887

Alignment Scores:
Pred. No.:      4.99
Score:          50.00
Percent Similarity: 86.7%
Best Local Similarity: 53.3%
Query Match:    54.3%
DB:             2

Length: 334
Matches: 8
Conservative: 5
Mismatch: 2
Indels: 0
Gaps: 0

US_10-664-025-43_COPY_53_103 (1-51) x HG9148 (1-334)

QY      4  GCCTTAGACTTTGANGCTCCGCTGCTTCATTGGACCGGT 48
|||||  :||:||||:||||:||||:||||:||||:
Db      11  AlaLeuSerIleLeuLeuIleLeuAlaLeuIleIleTrpMetGly 250

RESULT 2
SI7804

```

hypothetical protein 274 - Rhodobacter capsulatus

C;Species: Rhodobacter capsulatus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S17804
R;Burke, D.H.; Alberti, M.; Armstrong, G.A.; Hearst, J.E.
submitted to the EMBL Data Library, November 1991

A;Description: The complete nucleotide sequence of the 46 kb photosynthesis gene cluster
A;Reference number: S17803
A;Accession: S17804
A;Molecule type: DNA
A;Residues: 1-274 <EMB>
A;Cross-references: UNIPROT:P26158; UNIPARC:UPI0000013BE4E; EMBL:Z11165; NID:g46097; PIDN:

Alignment Scores:
Pred. No.: 16.1 Length: 274
Score: 47.00 Matches: 8
Percent Similarity: 78.6% Conservative: 3
Best Local Similarity: 57.1% Mismatches: 3
Query Match: 51.1% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x S17804 (1-274)

QY 7 TTGAAGCTTGATGCTCCCGTCGTCTTGGATTGGACCGGT 48
:: |||||
Db 123 ILSerAlaLeuIleLeuCySAlaValLeuVairThrGly 136

RESULT 3

A36353
DNA repair protein XRCC1 - human
C;Species: Homo sapiens (man)
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 05-Oct-2004
C;Accession: A36353
R;Thompson, L.H.; Brookman, K.W.; Jones, N.J.; Allen, S.A.; Carrano, A.V.
Mol. Cell. Biol. 10, 6160-6171, 1990

A;Title: Molecular cloning of the human XRCC1 gene, which corrects defective DNA strand
A;Reference number: A36353; MUID:91061722; PMID:2247054
A;Accession: A36353
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-633 <THO>
A;Cross-references: UNIPROT:P18887; UNIPARC:UPI00000139020; GB:M36089; NID:g340396; PIDN:
C;Genetics:
A;Gene: GDB:XRCCL1; RCC
A;Cross-references: GDB:I20737; OMIM:194360
A;Map position: 19q13.2-19q13.2
C;Superfamily: DNA-repair protein XRCC1

Alignment Scores:
Pred. No.: 14.5 Length: 633
Score: 47.00 Matches: 9
Percent Similarity: 76.9% Conservative: 1
Best Local Similarity: 69.2% Mismatches: 3
Query Match: 51.1% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x A36353 (1-633)

QY 51 AGCACCGGTCCAATCAAGACAGCAGGAGCATCAAGAC 13
|||||
Db 518 SerThrAspGluAnThrAspSerGluGluHisGlnGlu 530

RESULT 4

Tl3458
hypothetical protein Tl3458.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: Tl3458
R;Bavan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; New
submitted to the Protein Sequence Database, July 1999
A;Reference number: Z17587
A;Accession: Tl3458
A;Molecule type: DNA

A;Residues: 1-710 <BEV>
A;Cross-references: UNIPROT:Q9SU46; UNIPARC:UPI000003FF34; EMBL:AL109619; GSPDB:GN000062;
A;Experimental source: cultivar Columbia; BAC clone T19F6
C;Genetics:
A;Gene: ATSP:T19F6.90
A;Map position: 4
A;Introns: 73/3; 158/3; 200/2; 215/2; 242/3; 265/3; 329/3; 351/3; 374/3; 467/2; 508/1; 509/1; 510/1; 511/1; 512/1; 513/1; 514/1; 515/1; 516/1; 517/1; 518/1; 519/1; 520/1; 521/1; 522/1; 523/1; 524/1; 525/1; 526/1; 527/1; 528/1; 529/1; 530/1; 531/1; 532/1; 533/1; 534/1; 535/1; 536/1; 537/1; 538/1; 539/1; 540/1; 541/1; 542/1; 543/1; 544/1; 545/1; 546/1; 547/1; 548/1; 549/1; 550/1; 551/1; 552/1; 553/1; 554/1; 555/1; 556/1; 557/1; 558/1; 559/1; 560/1; 561/1; 562/1; 563/1; 564/1; 565/1; 566/1; 567/1; 568/1; 569/1; 570/1; 571/1; 572/1; 573/1; 574/1; 575/1; 576/1; 577/1; 578/1; 579/1; 580/1; 581/1; 582/1; 583/1; 584/1; 585/1; 586/1; 587/1; 588/1; 589/1; 590/1; 591/1; 592/1; 593/1; 594/1; 595/1; 596/1; 597/1; 598/1; 599/1; 600/1; 601/1; 602/1; 603/1; 604/1; 605/1; 606/1; 607/1; 608/1; 609/1; 610/1; 611/1; 612/1; 613/1; 614/1; 615/1; 616/1; 617/1; 618/1; 619/1; 620/1; 621/1; 622/1; 623/1; 624/1; 625/1; 626/1; 627/1; 628/1; 629/1; 630/1; 631/1; 632/1; 633/1; 634/1; 635/1; 636/1; 637/1; 638/1; 639/1; 640/1; 641/1; 642/1; 643/1; 644/1; 645/1; 646/1; 647/1; 648/1; 649/1; 650/1; 651/1; 652/1; 653/1; 654/1; 655/1; 656/1; 657/1; 658/1; 659/1; 660/1; 661/1; 662/1; 663/1; 664/1; 665/1; 666/1; 667/1; 668/1; 669/1; 670/1; 671/1; 672/1; 673/1; 674/1; 675/1; 676/1; 677/1; 678/1; 679/1; 680/1; 681/1; 682/1; 683/1; 684/1; 685/1; 686/1; 687/1; 688/1; 689/1; 690/1; 691/1; 692/1; 693/1; 694/1; 695/1; 696/1; 697/1; 698/1; 699/1; 700/1; 701/1; 702/1; 703/1; 704/1; 705/1; 706/1; 707/1; 708/1; 709/1; 710/1; 711/1; 712/1; 713/1; 714/1; 715/1; 716/1; 717/1; 718/1; 719/1; 720/1; 721/1; 722/1; 723/1; 724/1; 725/1; 726/1; 727/1; 728/1; 729/1; 730/1; 731/1; 732/1; 733/1; 734/1; 735/1; 736/1; 737/1; 738/1; 739/1; 740/1; 741/1; 742/1; 743/1; 744/1; 745/1; 746/1; 747/1; 748/1; 749/1; 750/1; 751/1; 752/1; 753/1; 754/1; 755/1; 756/1; 757/1; 758/1; 759/1; 760/1; 761/1; 762/1; 763/1; 764/1; 765/1; 766/1; 767/1; 768/1; 769/1; 770/1; 771/1; 772/1; 773/1; 774/1; 775/1; 776/1; 777/1; 778/1; 779/1; 780/1; 781/1; 782/1; 783/1; 784/1; 785/1; 786/1; 787/1; 788/1; 789/1; 790/1; 791/1; 792/1; 793/1; 794/1; 795/1; 796/1; 797/1; 798/1; 799/1; 800/1; 801/1; 802/1; 803/1; 804/1; 805/1; 806/1; 807/1; 808/1; 809/1; 810/1; 811/1; 812/1; 813/1; 814/1; 815/1; 816/1; 817/1; 818/1; 819/1; 820/1; 821/1; 822/1; 823/1; 824/1; 825/1; 826/1; 827/1; 828/1; 829/1; 830/1; 831/1; 832/1; 833/1; 834/1; 835/1; 836/1; 837/1; 838/1; 839/1; 840/1; 841/1; 842/1; 843/1; 844/1; 845/1; 846/1; 847/1; 848/1; 849/1; 850/1; 851/1; 852/1; 853/1; 854/1; 855/1; 856/1; 857/1; 858/1; 859/1; 860/1; 861/1; 862/1; 863/1; 864/1; 865/1; 866/1; 867/1; 868/1; 869/1; 870/1; 871/1; 872/1; 873/1; 874/1; 875/1; 876/1; 877/1; 878/1; 879/1; 880/1; 881/1; 882/1; 883/1; 884/1; 885/1; 886/1; 887/1; 888/1; 889/1; 890/1; 891/1; 892/1; 893/1; 894/1; 895/1; 896/1; 897/1; 898/1; 899/1; 900/1; 901/1; 902/1; 903/1; 904/1; 905/1; 906/1; 907/1; 908/1; 909/1; 910/1; 911/1; 912/1; 913/1; 914/1; 915/1; 916/1; 917/1; 918/1; 919/1; 920/1; 921/1; 922/1; 923/1; 924/1; 925/1; 926/1; 927/1; 928/1; 929/1; 930/1; 931/1; 932/1; 933/1; 934/1; 935/1; 936/1; 937/1; 938/1; 939/1; 940/1; 941/1; 942/1; 943/1; 944/1; 945/1; 946/1; 947/1; 948/1; 949/1; 950/1; 951/1; 952/1; 953/1; 954/1; 955/1; 956/1; 957/1; 958/1; 959/1; 960/1; 961/1; 962/1; 963/1; 964/1; 965/1; 966/1; 967/1; 968/1; 969/1; 970/1; 971/1; 972/1; 973/1; 974/1; 975/1; 976/1; 977/1; 978/1; 979/1; 980/1; 981/1; 982/1; 983/1; 984/1; 985/1; 986/1; 987/1; 988/1; 989/1; 990/1; 991/1; 992/1; 993/1; 994/1; 995/1; 996/1; 997/1; 998/1; 999/1; 1000/1; 1001/1; 1002/1; 1003/1; 1004/1; 1005/1; 1006/1; 1007/1; 1008/1; 1009/1; 1010/1; 1011/1; 1012/1; 1013/1; 1014/1; 1015/1; 1016/1; 1017/1; 1018/1; 1019/1; 1020/1; 1021/1; 1022/1; 1023/1; 1024/1; 1025/1; 1026/1; 1027/1; 1028/1; 1029/1; 1030/1; 1031/1; 1032/1; 1033/1; 1034/1; 1035/1; 1036/1; 1037/1; 1038/1; 1039/1; 1040/1; 1041/1; 1042/1; 1043/1; 1044/1; 1045/1; 1046/1; 1047/1; 1048/1; 1049/1; 1050/1; 1051/1; 1052/1; 1053/1; 1054/1; 1055/1

RESULT 9
A83118
probable ATP-binding component of ABC transporter PA4223 [imported] - *Pseudomonas aeruginosa*
C:Species: *Pseudomonas aeruginosa*

C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C; Accession: A83118
R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A; Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A; Reference number: A82950; MUID:20437337; PMID:10984043
A; Accession: A83118
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-570 <S>O>
A; Cross-references: UNIPROT:Q9HWG6; UNIPARC:UPI000000C5EB7; GB:AE0040839; GB:AE004091; NID
A; Experimental source: strain PA01
C; Genetics:
A; Gene: PA4223

Alignment Scores:					
Pred.	No.:		Length:		570
Score:		31.5	Matches:		10
Percent Similarity:		45.00	Conservative:		1
Best Local Similarity:		73.3%	Mismatches:		4
Query Match:		66.7%	Indels:		0
DB:		2	Gaps:		0
 US-10-564-025-43_COPY_53_103 (1-51) x A83118 (1-570)					
Qy	4	CGGTTAGAACTTGTGATGCCTCGCTGTCITTGATTGGACCGGT	48		
Db	52	AlaLeuLeuAlaLeuValLeuLeuAlaValLeuAlaTrpLeuGly	66		

RESULT 10

D69815 conserved hypothetical protein yfn1 - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
C/Accession: D69815
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertain
C.; Brun, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabet, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Xunano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, C.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Toso, V.; Uchiyama,
t.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D07613
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-653 <KUN>
A:Cross-references: UNIPROT:O06487; UNIPARC:UPI000006006F; GB:Z99107; GB:AL009126; NID:9
A:Experimental source: strain 168
C:Genetics:
A:Gene: yfml
C:Superfamily: membrane sulfatase, H11245 type

Alignment Scores:		
Pred. No.:	31	Length: 653
Score:	45.00	Matches: 7
Percent Similarity:	100.0%	Conservative: 5
Best Local Similarity:	58.3%	Mismatches: 0
Query Match:	48.9%	Indels: 0
DB:	2	Gaps: 0

US-10-664-025-43 COPY 53 103 (1-51) x D69815 (1-653)

RESULT 13

T10756
Nel-homolog protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10756
R:Kuroda, S.; Tokunaga, C.; Kiyohara, Y.; Konishi, H.; Matsushashi, S.; Kikkawa, U.
submitted to the EMBL Data Library, November 1998
A:Description: Protein kinase C-binding protein.
A:Reference number: Z17122
A:Accession: T10756
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-810 <KUR>
A:Cross-references: UNIPROT:Q62919; UNIPARC:UPI000012PF2B; EMBL:U48246; NID:G3851179; P
A:Experimental source: strain Sprague-Dawley, brain

Alignment Scores:

Pred. No.:	44.1	Length:	810
Score:	44.00	Matches:	7
Percent Similarity:	56.2%	Conservative:	2
Best Local Similarity:	43.8%	Mismatches:	7
Query Match:	47.8%	Indels:	0
DB:	2	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x T10756 (1-810)

QY 3 GGCGTTAGAGTCTTGATGCTCTCGCTCTTGATTTGGACCGGTGC 50
Db 585 GlyThrTyrSerLeuSerGlyGluSerCysIleAspGluCys 600

RESULT 14

FECCLCE

ferredoxin 2[4Fe-4S] - Clostridium sp.

C:Species: Clostridium sp.

C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 12-Jul-2004

C:Accession: A00200

R/Tanaka, M.; Haniu, M.; Yasunobu, K.T.; Jones, J.B.; Stadtman, T.C.

Biochemistry 13, 5284-5289, 1974

A:Title: Amino acid sequence determination of the Clostridium M-E ferredoxin and a com

A:Reference number: A00200; MUID:75054829; PMID:4433520

A:Molecule type: protein

A:Residues: 1-55 <TAN>

A:Cross-references: UNIPARC:UPI000012A6B3

A:Experimental source: strain M-E

C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein

F;1-55/Domain: ferredoxin 2[4Fe-4S] homology <FER>

F;8,11,14,47/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

F;18,37,40,43/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Alignment Scores:

Pred. No.:	90.7	Length:	55
Score:	43.00	Matches:	7
Percent Similarity:	66.7%	Conservative:	3
Best Local Similarity:	46.7%	Mismatches:	5
Query Match:	46.7%	Indels:	0
DB:	1	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x FECCLCE (1-55)

QY 6 GTTAGAGTCTTGATGCTCTCGCTCTTGATTTGGACCGGTGC 50
Db 29 ValArgValIleAspAlaAspLysCysIleAspCysGlyAlaCys 43

RESULT 15

T01671

vpu protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 20-Jun-2000

C:Accession: T01671

R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.

Cell 46, 63-74, 1986

A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two is
A:Reference number: Z14389; MUID:86245056; PMID:2424612
A:Accession: T01671
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-81 <ALI>
A:Cross-references: UNIPARC:UPI000017865F; EMBL:K03456; NID:G328018
C:Superfamily: HIV-1 vpu protein

Alignment Scores:

Pred. No.:	86.4	Length:	81
Score:	43.00	Matches:	6
Percent Similarity:	86.7%	Conservative:	7
Best Local Similarity:	40.0%	Mismatches:	2
Query Match:	46.7%	Indels:	0
DB:	2	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x T01671 (1-81)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTCTTGATTTGGACC 45
Db 10 ValAlaLeuValValThrLeuIleIleAlaIleValValThr 24

Search completed: May 9, 2006, 10:10:07

Job time : 5.4 secs

This Page Blank (uspto)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:02:52 ; Search time 2.62 Seconds

(without alignments)
4120.070 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_103

Perfect score: 92

Sequence: 1 atggcgcttagaagctctgat.....tcttgattggaccggtgct 51

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+np.model -DRV=xlh
-Q=/abss/ABSSWEB_spool/US10664025/runat_08052006_173450_28308/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNIT8-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss07
-USER=US10664025 @CGN_1_1_266 @runat_08052006_173450_28308 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result N°	Score	Query Match	Length	ID	Description
1	81	88.0	158	Q8N9U6 HUMAN	Q8N9U6 homo sapien
2	81	88.0	158	Q86WS3 HUMAN	Q86WS3 homo sapien
3	62	67.4	164	Q4FZG8 MOUSE	Q4FZG8 mus musculus
4	53	57.6	174	Q8T8D7 TAESA	Q8T8D7 taenia sagi
5	53	57.6	432	Q5N0H7_SYNP6	Q5N0H7 synchococc
6	51	55.4	132	Q5CN74_CRYHO	Q5CN74 cryptospori
7	50	54.3	79	Q29172_PIG	Q29172 sus scrofa
8	50	54.3	334	Y378_METTH	Q26478 methanobact
9	50	54.3	507	Q4WR18_ASPFU	Q4WR18 aspergillus
10	49.5	53.8	1152	Q5AS17_EMENI	Q5AS17 aspergillus
11	49	53.3	674	Q5AN72_DICDI	Q5AN72 dictyosteli
12	49	53.3	1098	Q6RR14_9BROM	Q6RR14 parietaria
13	48	52.2	81	Q8AE43_9HIV1	Q8AE43 human immun
14	47	51.1	81	Q8AD88_9HIV1	Q8AD88 human immun
15	47	51.1	199	Q94MU5_9CAUD	Q94MU5 bacterioph
16	47	51.1	219	Q6SNL4_CEBAP	Q6SNL4 cebus apell

17	47	51.1	223	2	Q6LSD9_PHOPR	Q6LSD9 photobacter
18	47	51.1	250	2	Q9XB60_ERWCA	Q9XB60 erwinia car
19	47	51.1	274	1	YPU2_RHOCA	P26158 rhodobacter
20	47	51.1	306	2	Q4S8A4_TETNG	Q4S8A4 tetraodon n
21	47	51.1	440	2	Q72WP7_DESVH	Q72WP7 desulfovibr
22	47	51.1	633	1	XRCC1_HUMAN	P18887 homo sapien
23	47	51.1	633	2	Q6IBS4_HUMAN	Q6IBS4 homo sapien
24	47	51.1	647	2	Q59HH7_HUMAN	Q59HH7 homo sapien
25	47	51.1	857	2	Q4WQ23_ASPFU	Q4WQ23 aspergillus
26	47	51.1	1333	2	Q4PIN0_USTMA	Q4PIN0 ustilago ma
27	46	50.0	19	2	Q90SE4_9HIV1	Q90SE4 human immun
28	46	50.0	78	2	Q8AD96_9HIV1	Q8AD96 human immun
29	46	50.0	80	2	Q4PU48_9HIV1	Q4PU48 human immun
30	46	50.0	80	2	Q4PU58_9HIV1	Q4PU58 human immun
31	46	50.0	81	2	Q5G7F2_9HIV1	Q5G7F2 human immun
32	46	50.0	81	2	Q8ADY6_9HIV1	Q8ADY6 human immun
33	46	50.0	82	2	Q7SP88_9HIV1	Q7SP88 human immun
34	46	50.0	99	2	P88153_9HIV1	P88153 human immun
35	46	50.0	101	2	Q8UPP1_9HIV1	Q8UPP1 human immun
36	46	50.0	109	2	Q6EQU5_ORYSA	Q6EQU5 oryza sativ
37	46	50.0	230	2	Q8G6X3_BIFLO	Q8G6X3 bifidobacte
38	46	50.0	320	2	Q4SQC4_TETNG	Q4SQC4 tetraodon n
39	46	50.0	354	2	Q6GJF0_XENTR	Q6GJF0 xenopus tro
40	46	50.0	572	2	Q22981_ARATH	Q22981 arabidopsis
41	46	50.0	614	2	Q4SMK4_TETNG	Q4SMK4 tetraodon n
42	46	50.0	654	2	Q4Q635_LEIMA	Q4Q635 leishmania
43	46	50.0	694	2	Q82FLI_STRAW	Q82FLI streptomyce
44	46	50.0	709	2	Q8RWU9_ARATH	Q8RWU9 arabidopsis
45	46	50.0	710	2	Q9SU46_ARATH	Q9SU46 arabidopsis

ALIGNMENTS

RESULT 1

Q8N9U6 HUMAN
ID Q8N9U6_HUMAN PRELIMINARY; PRT; 158 AA.
AC Q8N9U6_
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ36198.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Tagiya S., Komai F., Hara R., Oshima A., Sasaki N.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

RA Mateumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Iada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK093517; BAC04191.1; -; mRNA.
 DR Ensembl; ENSG00000149507; Homo sapiens.
 SQ SEQUENCE 158 AA; 17942 MW; EE943D70A8A391E8 CRC64;

Alignment Scores:
 Pred. No.: 2.75e-05 Length: 158
 Score: 81.00 Matches: 17
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 88.0% Indels: 0
 DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q8N9U6_HUMAN (1-158)

QY 1 ATGCGGTAGAGTCTTCATGCTCTCGTCTGTTGATTGGACCGGTGCT 51
 Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 2

Q86WS3 HUMAN
 ID Q86WS3; Q86WS3; PRT; 158 AA.
 AC Q86WS3;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
 DE Hypothetical protein FLJ36198.
 GN Name=FLJ36198;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RG NIH MGC Project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RG Director MGC Project;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC048121; AAH48121.1; -; mRNA.
 DR EMBL; BC036256; AAH36256.1; -; mRNA.
 DR Ensembl; ENSG00000149507; Homo sapiens.
 KW Hypothetical protein.
 SQ SEQUENCE 158 AA; 17971 MW; EE843D6FB8AB81FF CRC64;

Alignment Scores:
 Pred. No.: 2.75e-05 Length: 158
 Score: 81.00 Matches: 17
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 88.0% Indels: 0
 DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q86WS3_HUMAN (1-158)

QY 1 ATGCGGTAGAGTCTTCATGCTCTCGTCTGTTGATTGGACCGGTGCT 51
 Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAla 17

RESULT 3

Q4FZG8 MOUSE
 ID Q4FZG8; Q4FZG8; PRT; 164 AA.
 AC Q4FZG8;
 DT 13-SEP-2005 (TREMBLrel. 31, Created)
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Oocytes;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Oocytes;
 RG NIH MGC Project;
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC099498; AAH99498.1; -; mRNA.
 KW Hypothetical protein.
 SQ SEQUENCE 164 AA; 18849 MW; 49ADE19216BE8606 CRC64;

Alignment Scores:

Pred. No.: 0.0778 Length: 164
 Score: 62.00 Matches: 13
 Percent Similarity: 88.2% Conservative: 2
 Best Local Similarity: 76.5% Mismatches: 2
 Query Match: 67.4% Indels: 0


```
DB:
US-10-664-025-43_COPY_53_103 (1-51) x Q4FZG8_MOUSE (1-164)
Gaps: 0

QY 1 ATGCGGTAGAGTCTTGATGCTCCGCTGCTGTGATTTGGACCGGTGCT 51
|||||
Db 1 MetAlaLeuGluValLeuValTyrLeuAlaValLeuValTrpThrCysAla 17
|||||

RESULT 4
Q8T8D7_TAESA
ID Q8T8D7_TAESA PRELIMINARY; PRT; 174 AA.
AC Q8T8D7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical protein (Fragment).
OS Taenia saginata (Beef tapeworm).
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Taenia.
OX NCBI_TaxID=6206;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Kenyan;
RA Garate T.;
RL EMBL: AJ430566; CAD23242.1; -; mRNA.
DR InterPro: IPR003961; FN_III.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS00853; FN3; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 174 AA; 18702 MW; 645D04766AC068AF CRC64;

Alignment Scores:
Pred. No.: 3.37 Length: 174
Score: 53.00 Matches: 9
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 75.0% Mismatches: 0
Query Match: 57.6% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q8T8D7_TAESA (1-174)
QY 16 TTGATGCTCCTCGCTCTTGATTTGGACCGGTGCT 51
|||||
Db 4 LeuMetLeuAlaLeuValLeuValTrpThrGlyser 15
|||||

RESULT 5
Q5NOH7_SYNP6
ID Q5NOH7_SYNP6 PRELIMINARY; PRT; 432 AA.
AC Q5NOH7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Homoserine dehydrogenase.
GN Name: thrA; OrderedLocusNames=syc2003_C;
OC Synecococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=269084;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCC6301;
RA Sugita M.;
RT "Complete genome structure of the unicellular cyanobacterium Anacystis
nidulans 6301 (Synecococcus sp. PCC6301).";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF008231; BAD80193.1; -; Genomic DNA.
DR GO: GO:0003984; Fructose 1,6-bisphosphate synthase activity; IEA.
DR GO: GO:0016597; F-1-AMP-dependent pyrophosphatase activity; IEA.
DR GO: GO:0004412; F-homoserine dehydrogenase activity; IEA.
DR GO: GO:0009082; P-branching chain family amino acid biosynthesis; IEA.
DR GO: GO:0008152; P-metabolism; IEA.
DR InterPro: IPR004789; Acolac_synthem.
```

```
DR InterPro: IPR002912; ACT.
DR InterPro: IPR005106; Hmer_dh_NAD_bind.
DR InterPro: IPR001342; Homoser_dh_hydrog.
DR Pfam: PF01842; ACT; 1.
DR Pfam: PF00742; Homoserine dh; 1.
DR Pfam: PF03447; NAD binding_3; 1.
DR ProDom: PD002844; Acolac_synthem; 1.
DR PROSITE; PS01042; HOMOSER_DHGENASE; 1.
KW Complete proteome.
SQ SEQUENCE 432 AA; 45628 MW; 3FE3A02F84FD75B4 CRC64;

Alignment Scores:
Pred. No.: 3.54 Length: 432
Score: 53.00 Matches: 10
Percent Similarity: 80.0% Conservative: 2
Best Local Similarity: 66.7% Mismatches: 3
Query Match: 57.6% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q5NOH7_SYNP6 (1-432)
QY 3 GCGTTAGAGTCTTGATGCTCCTCGCTCTTGATTTGGACCGG 47
|||||
Db 40 GlyValArgSerLeuAspLysProArgSerValAspIleSerArg 54
|||||

RESULT 6
Q5CN74_CRYHO
ID Q5CN74_CRYHO PRELIMINARY; PRT; 132 AA.
AC Q5CN74;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG4090-PA.
GN ORFNames=Chro.80611;
OC Cryptosporidium hominis.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=237895;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TU502;
RA Xu P.; Widmer G.; Wang Y.; Ozaki L.S.; Alves J.M.; Serrano M.G.;
RA Puiu D.; Manque P.; Akiyoshi D.; Mackey A.J.; Pearson W.R.; Dear P.H.;
RA Bankier A.T.; Peterson D.L.; Abrahamson M.S.; Kapur V.; Tsipori S.;
RA Buck G.A.;
RT "The genome of Cryptosporidium hominis.";
RL Nature 431:1107-1112(2004).
DR EMBL: AAELO1000015; EAL38051.1; -; Genomic DNA.
SQ SEQUENCE 132 AA; 14733 MW; 8721106746E1F5F CRC64;

Alignment Scores:
Pred. No.: 7.65 Length: 132
Score: 51.00 Matches: 9
Percent Similarity: 90.9% Conservative: 1
Best Local Similarity: 81.8% Mismatches: 1
Query Match: 55.4% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q5CN74_CRYHO (1-132)
QY 45 GGTCCAAATCAAGACGAGCAGGATCAAGAC 13
|||||
Db 72 GlyProAsnGlnAspSerGluSerAsnGlnAsp 82
|||||

RESULT 7
Q29172_PIG
ID Q29172_PIG PRELIMINARY; PRT; 79 AA.
AC Q29172;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Orf protein (Fragment).
GN Name=orf;
```

KW	Complete proteome; Hypothetical protein; Transmembrane.
FT	TRANSMEM 7 27 Potential.
FT	TRANSMEM 33 53 Potential.
FT	TRANSMEM 120 140 Potential.
FT	TRANSMEM 142 162 Potential.
FT	TRANSMEM 218 238 Potential.
FT	TRANSMEM 247 267 Potential.
FT	TRANSMEM 277 297 Potential.
FT	TRANSMEM 300 320 Potential.
SQ	SEQUENCE 334 AA; 36959 MW; 46D280EC683CB611 CRC64;
 Alignment Scores:	
Pred. No.:	12.2 Length: 334
Score:	50.00 Matches: 8
Percent Similarity:	86.7% Conservative: 5
Best Local Similarity:	53.3% Mismatches: 2
Query Match:	54.3% Indels: 0
DB:	1 Gaps: 0
 US-10-664-025-43_COPY_53_103 (1-51) x Y378_METTH (1-334)	
Qy	4 GCCTTAGAAGTCTTGATGCTCCTCGCTGTCTTGATTGCACGGT 48 : ::: ::: ::: :::
Dd	11 AlaleuSerlleuleuileuialeuialeullellelTrpMetGly 25 : ::: ::: ::: :::
 RESULT 9	
Q4WR18 ASPFU	ID Q4WR18 ASPFU PRELIMINARY; PRT; 507 AA.
AC	Q4WR18; 507 AA.
DC	13-SEP-2005 (TEMBLrel. 31, Created)
DT	13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DD	13-SEP-2005 (TEMBLrel. 31, Last annotation update)
DE	Cytochrome p450 monooxygenase, putative.
GN	O8FNames=Afu4g14790.
GE	Aspergillus fumigatus Af293.
OS	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX	NB1_TaxID=330879;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=Af293;
RA	Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA	Aroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA	Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA	Farnham M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA	Fosker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA	Goldman G.H., Gomi K., Griffith-Jones S., Guilliam R., Haas B.,
RA	Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA	Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulikarni R.,
RA	Kumagai T., Latfon A., Latge J.-P., Li W., Lord A., Lu C.,
RA	Majors T.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA	Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA	Penalva M.A., Pertea M., Price C., Pritchard B.L., Quail M.A.,
RA	Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,
RA	Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA	Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA	Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA	Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA	White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA	Machida M., Hall N., Barrrell B., Denning D.W.;
RT	"Genomic sequence of the pathogenic and allergenic filamentous fungus
RL	Aspergillus fumigatus";
RL	Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC	-!- CAUTION: The sequence shown here is derived from an
CC	EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC	preliminary data.
CC	-!- SIMILARITY: Belongs to the cytochrome P450 family.
CC	EMBL: AAAF01000005; EAL89316.1; -; Genomic_DNA.
DR	InterPro: IPRO01128; Cytochrome_p450.
DR	InterPro: IPRO02401; EP450I.
DR	InterPro: IPRO01865; Ribosomal_S2.
DR	Pfam: PF00067; p450; 1.
DR	PRINTS: PR00463; EP450I.

DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN 1.
DR PROSITE; PS00962; RIBOSOMAL_S2; UNKNOWN 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 507 AA; 57810 MW; 16FIDE6590612F32 CRC64;

Alignment Scores:		
Pred. No.:	12.5	Length:
Score:	50.00	Matches:
Percent Similarity:	86.7%	Conservative:
Best Local Similarity:	53.3%	Mismatches:
Query Match:	54.3%	Indels:
DB:	2	Gaps:
		507

US-10-664-025-43 COPY 53 103 (1-51) x Q4WR18 ASPFU (1-507)

QY

1 ATGGCGTTAGAAAGCTTGTGATGCTCCTCGTGTCCTGTGATTGGACC 45
||||||| : ::::: ||||||| : ::::: |||||

Dd

1 MetAlaLeuProIleLeuCysLeuAlaValIleLeuTrpThr 15

RESULT 10

Q5ASI7	EMENI	PRELIMINARY;	PRT;	1152	AA.
ID	Q5ASI7	EMENI	PRELIMINARY;		
AC	Q5ASI7;				
DT	10-WAY-2005	(TREMBLrel. 30, Created)			
DT	10-WAY-2005	(TREMBLrel. 30, Last sequence update)			
DT	10-WAY-2005	(TREMBLrel. 30, Last annotation update)			
DT	10-WAY-2005	(TREMBLrel. 30, Last annotation update)			
DE	Hypothetical protein.				
GN	ORFNames=AN8743.2;				
DN					
OS	Aspergillus nidulans FGSC A4.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiales; Trichocomaceae; Emericella.				
OX	NCBI_TaxID=227321;				
RN	[1]				

RP NUCLEOTIDE SEQUENCE.

STRAIN=FGSC A4;

RR Birren B., Nuebaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
RA Boukhgater B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Kartas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schnupback R., Seaman S., Savery P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;

"Genome Sequence of *Aspergillus nidulans*.";
Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.

-!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.

DR EMBL; AACD01000161; EAA60536.1; -; Genomic DNA.

Hypothetical protein. KW
SEQUENCE 1152 AA; 125816 MW; ACBEE3B1AFFECCDB CRC64;

Alignment Scores:

Pred. No.:	16.2	Length:	1152
Score:	49.50	Matches:	10
Percent Similarity:	66.7%	Conservative:	2
Best Local Similarity:	55.6%	Mismatches:	3
Query Match:	53.8%	Indels:	3

DB:	2	Gap8:	1
US-10-664-025-43	COPY 53	103 (1-51) x Q5AS17	ENENI (1-1152)

Qy 45 GGTCAAATCAAGAC-----AGCGAGGAGCATCAAGACTTCTAACGCCAT 1
Db 760 GlyProAsnArgAspAsnProHisSerGlnGluHisAspAspPheSerHisHis 7

RESULT 11

Q54N72_DICDI	PRELIMINARY;	PRT;	674 AA.
ID	Q54N72_DICDI		
AC	Q54N72;		
DT	13-SEP-2005 (TrEMBLrel. 31, Created)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)		
DE	Hypothetical protein.		
GN	ORFNames=DD80186513;		
OS	Dictyostelium discoideum (Slime mold).		
OC	Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium		
OX	NCBI_TaxID=44689;		
	[1]		
RP	NUCLEOTIDE SEQUENCE.		

RC STRAIN=AX4;

RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.
 RA Sungang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
 RA Tungal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
 RA Pilcher K., Chen G., Saunders D., Sodergren E., David P.,
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
 RA Farbrother P., Deasany B., Just E., Worio T., Roost R., Churcher C.,
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
 RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louldsegh H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
 RA Shaulsky G., Schreicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
 RA Williams J., Dear P.H., Nogel A.A., Barrell B., Kuspa A.;
 RT "The genome of the social amoeba Dictyostelium discoideum.;
 CC Nature 0:0-0(2005).
 CC
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC
 DR EMBL; AAFO1000124; EAI64570.1; -; Genomic_DNA.
 DR Hypothetical protein.
 SQ
 SQ SEQUENCE 674 AA; 77632 MW; 920A385423EC59BD CRC64;

Alignment Scores:

Pred. No.:	19.3	Length:	674
Score:	49.00	Matches:	9
Percent Similarity:	81.2%	Conservative:	4
Best Local Similarity:	56.2%	Mismatches:	3
Query Match:	53.3%	Indels:	0
DB:	2	Gaps:	0

US-10-664-025-43 COPY 53 103 (1-51) X 054N72 DICDI (1-674)

QY 49 CACCGTCCAANTCAAGACAGCGGAGCATCAAGACTTCTTAACGCCA 2
:::|||||
db 599 AsnArgSerLysSerArgGlnArgSerThrThrArgIleAlaSerPro 6

RESULT 12

RESULTS	Q6RK14_9BROM	Q6RK14_9BROM PRELIMINARY;	PRT; 1098 AA.
ID	Q6RK14_9BROM	Q6RK14_9BROM PRELIMINARY;	
AC	Q6RK14_9BROM	Q6RK14_9BROM PRELIMINARY;	
DT	05-JUL-2004	(TREMBlErel. 27, Created)	
DT	05-JUL-2004	(TREMBlErel. 27, Last sequence update)	
DT	05-JUL-2004	(TREMBlErel. 27, Last annotation update)	
DE	P1 protein.		
OS	Parietaria mottle virus.		

```
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Iarivirus.
OX NCBI_TaxID=64958;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Scott S.W., Zimmerman M.T., Rankin D.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL; AY496068; AAR8259.1; -; Genomic RNA.
DR GO; GO:0008174; F:RNA methyltransferase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR000606; Viral helicase.
DR InterPro; IPR002588; V_methyltrans.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Pfam; PF01660; Vmethyltrans1; 1.
SQ SEQUENCE 1098 AA; 124146 MW; 4FA11DC1D83D1347 CRC64;

Alignment Scores:
Pred. No.: 19.9 Length: 1098
Score: 49.00 Matches: 8
Percent Similarity: 73.3% Conservative: 3
Best Local Similarity: 53.3% Mismatches: 4
Query Match: 53.3% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q6RK14_9BROM (1-1098)

QY 6 GTTAGAGCTTTGATGCTCTCGTCTGTTGATTTGGACCGGTGC 50
|||||:|||||: ||| |||||||: |||
571 ValArgAenLeuGluPheProSerCysLeuAspValSerGluCys 585

RESULT 13
Q8AE43_9HIV1
ID Q8AE43_9HIV1 PRELIMINARY; PRT; 81 AA.
AC Q8AE43;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vpu protein.
GN Name=vpu;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus;
OC Primate lentivirus group.
OX NCBI_TaxID=11676;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGB25647;
RX MEDLINE=22375625; PubMed=12487816; DOI=10.1089/089922202320886325;
RA Harris M.E., Serwadda D., Sewankambo N., Kim B., Kigozi G.,
RA Kiwanuka N., Phillips J.B., Wabwire F., Meehen M., Lutalo T.,
RA Lane J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 near full length HIV type 1 genome sequences from Rakai
RT District, Uganda, subtype D and AD recombinants predominate.";
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGB25647;
RA Harris M.E., Birx D.L., Robb M.L.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGB25647;
RA Kim B., Phillips J.B., Lane J.R., Merling R., McCutchan F.E.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGB25647;
RA Lutalo T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RP NUCLEOTIDE SEQUENCE.
RN [5]
RP Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
```

```
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Meehen M., Wawer M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGB25647;
RA Serwadda S., Sewankambo N., Wabwire F., Kigozi G., Kiwanuka N.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF484481; AAN73479.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019076; P:viral release; IEA.
DR InterPro; IPR008187; Vpu.
DR InterPro; IPR009032; Vpu_cyt.
DR Pfam; PF00558; Vpu; 1.
SQ SEQUENCE 81 AA; 9322 MW; 98B7887582935D15 CRC64;

Alignment Scores:
Pred. No.: 26.1 Length: 81
Score: 48.00 Matches: 8
Percent Similarity: 86.7% Conservative: 5
Best Local Similarity: 53.3% Mismatches: 2
Query Match: 52.2% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_103 (1-51) x Q8AE43_9HIV1 (1-81)

QY 1 ATGGCGTTAGAGCTTGTGATGCTCTCGTCTGTTGATTTGGACC 45
|||||:|||||: ||| |||||||: |||
10 ValAlaLeuValAlaLeuLeuAlaLeuLeuAlaLeuValTrpThr 24

RESULT 14
Q8AD88_9HIV1
ID Q8AD88_9HIV1 PRELIMINARY; PRT; 81 AA.
AC Q8AD88;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vpu protein.
GN Name=vpu;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus;
OC Primate lentivirus group.
OX NCBI_TaxID=11676;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGK38855;
RX MEDLINE=22375625; PubMed=12487816; DOI=10.1089/089922202320886325;
RA Harris M.E., Serwadda D., Sewankambo N., Kim B., Kigozi G.,
RA Kiwanuka N., Phillips J.B., Wabwire F., Meehen M., Lutalo T.,
RA Lane J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 near full length HIV type 1 genome sequences from Rakai
RT District, Uganda, subtype D and AD recombinants predominate.";
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGK38855;
RA Harris M.E., Birx D.L., Robb M.L.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGK38855;
RA Kim B., Phillips J.B., Lane J.R., Merling R., McCutchan F.E.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=99UGK38855;
RA Lutalo T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
```

RC STRAIN=99UGK38855;
RA Meehen M., Wawer M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99UGK38855;
RA Serwadda S., Sewankambo N., Wabwire F., Kigozi G., Kiwanuka N.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF484520; AAN73822.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019076; P:viral release; IEA.
DR InterPro; IPR008187; Vpu.
DR InterPro; IPR009032; Vpu_cyt.
DR Pfam; PF00558; Vpu; 1.
SQ SEQUENCE 81 AA; 9317 MW; 430257630240FDEC CRC64;

Alignment Scores:

Pred. No.:	39.7	Length:	81
Score:	47.00	Matches:	8
Percent Similarity:	86.7%	Conservative:	5
Best Local Similarity:	53.3%	Mismatches:	2
Query Match:	51.1%	Indels:	0
DB:	2	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x Q8AD88_9HIV1 (1-81)

QY 1 ATGGCGTTAGAGCTTGTGATGCTCTCGCTGCTGTGATTGGACC 45
Db 10 ValAlaLeuValValAlaLeuLeuLeuAlaileValValTrpThr 24

RESULT 15

Q94MU5_9CAUD
ID Q34MU5_9CAUD PRELIMINARY; PRT; 199 AA.
AC Q94MU5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P24.
OS Bacteriophage Mx8.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses.
OX NCBI_TaxID=49964;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Youderian P., Walthers D., Salmi D., Magrini V., Hartzell P.L.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF396866; AAK94359.1; -; Genomic DNA.
SQ SEQUENCE 199 AA; 22023 MW; F893907D28FD4DD7 CRC64;

Alignment Scores:

Pred. No.:	41.7	Length:	199
Score:	47.00	Matches:	10
Percent Similarity:	68.8%	Conservative:	1
Best Local Similarity:	62.5%	Mismatches:	5
Query Match:	51.1%	Indels:	0
DB:	2	Gaps:	0

US-10-664-025-43_COPY_53_103 (1-51) x Q94MU5_9CAUD (1-199)

QY 49 CACCGGTCCAAATCAAGCAGCAGGAGCATCAAGACTTCTAACGCCA 2
Db 26 HisArgGlnLysSerArgGlnArgAspAlaThrArgArgTrpPro 41

Search completed: May 9, 2006, 10:09:14
Job time : 16.1 secs

This Page Blank (uspro)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 14:08:32 ; Search time 1497.4 Seconds

(without alignments)
7460.154 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511

Perfect score: 459

Sequence: 1 atggcgcttagaagctctgat.....agttaggattataatcttc 459

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	459	100.0	512	BD107959	BD107959 EST and e
2	459	100.0	512	AR414950	AR414950 Sequence
3	459	100.0	512	AX969240	AX969240 Sequence
4	453	98.7	477	CQ737424	CQ737424 Sequence
5	453	98.7	642	BD109278	BD109278 EST and e
6	453	98.7	642	AR413725	AR413725 Sequence
7	453	98.7	642	AX970559	AX970559 Sequence
8	453	98.7	691	CS072271	CS072271 Sequence
9	453	98.7	1337	AX574405	AX574405 Sequence
10	453	98.7	1480	AX600204	AX600204 Sequence
11	453	98.7	1502	BC036256	BC036256 Homo sapi
12	453	98.7	1646	BC048121	BC048121 Homo sapi
13	453	98.7	1898	BD275948	BD275948 62 Human
14	451.4	98.3	1492	AX748179	AX748179 Sequence
15	451.4	98.3	1492	AX093517	AX093517 Homo sapi
16	446	97.2	497	AX588689	AX588689 Sequence
17	429	93.5	500	AX588688	AX588688 Sequence
18	341.6	74.4	560	BD110503	BD110503 EST and e

19	341.6	74.4	560	6	AR414950
20	341.6	74.4	560	6	AX971784
21	272	59.3	372	6	BD119756
22	272	59.3	372	6	AR424203
23	272	59.3	372	6	AX984897
24	246	53.6	469	6	BD109279
25	246	53.6	469	6	AR413726
26	246	53.6	469	6	AX970550
27	246	53.6	470	6	CS072276
28	234.6	51.1	655	9	BC099498
29	183	39.9	10115	6	AX574404
30	183	39.9	100317	14	AP000409
31	183	39.9	101031	14	AP000622
32	183	39.9	176944	8	AP000790
33	183	39.9	188367	8	AC013807
34	101.2	22.0	139212	14	AC154077
35	92.2	20.1	132494	9	AC126036
36	92.2	20.1	203047	9	AC127289
37	92.2	20.1	245162	14	AC110632
38	90.2	19.7	784	6	BD125165
39	90.2	19.7	784	6	BD126983
40	90.2	19.7	784	6	CQ780456
41	90.2	19.7	784	6	CQ782274
42	90.2	19.7	1067	8	BC066327
43	90.2	19.7	1126	6	BD127476
44	90.2	19.7	1126	6	CQ783102
45	90.2	19.7	1126	8	AK075086

ALIGNMENTS

RESULT 1

BD107959

LOCUS

BD107959 EST and encoded human protein.

DEFINITION

BD107959

ACCESSION

BD107959.1

VERSION

JP 2002010789-A/36

KEYWORDS

SOURCE

Homo sapiens

ORGANISM

REFERENCE

1 (bases 1 to 512)

AUTHORS

Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.

TITLE

EST and encoded human protein

JOURNAL

Patent: JP 2002010789-A 36 15-JAN-2002;

GENSET CORP

COMMENT

OS Homo sapiens (human)

PN JP 2002010789-A/36

PD 15-JAN-2002

PF 07-AUG-2000 JP 2000280989

PR 05-AUG-1999 US 60/147499

PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI

GIORDANO

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

C12N1/21,

PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC

C12N15/00

CC Von Heijne matrix

CC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

CC score 10.6999998092651

CC seq VMLLI AVLITGA/EN

FT CDS

FT sig_peptide 53..511

Location/Qualifiers

1..512

Location/Qualifiers

source

ORIGIN

BD107959 512 bp DNA linear PAT 18-SEP-2002

EST and encoded human protein.

BD107959

BD107959.1

GI:23202777

JP 2002010789-A/36

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

1 (bases 1 to 512)

Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.

EST and encoded human protein

Patent: JP 2002010789-A 36 15-JAN-2002;

GENSET CORP

OS Homo sapiens (human)

PN JP 2002010789-A/36

PD 15-JAN-2002

PF 07-AUG-2000 JP 2000280989

PR 05-AUG-1999 US 60/147499

PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI

GIORDANO

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

C12N1/21,

PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC

C12N15/00

CC Von Heijne matrix

CC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

CC score 10.6999998092651

CC seq VMLLI AVLITGA/EN

FT CDS

FT sig_peptide 53..511

Location/Qualifiers

1..512

Location/Qualifiers

source

ORIGIN

181 CGGATACATACATATGATATGATTTATATATCTTGTTCGTCATGTCGTCATCAGGACA 240
233 CGGATACATACATATGATATGATTTATATATCTTGTTCGTCATGTCGTCATCAGGACA 292
241 AGGGTAGTTCTTCAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
293 AGGGTAGTTCTTCAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
301 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
353 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 412
361 CTTCACACCACTTCTACTGAGATGAATAAATTTGATCTAGTCTCTTTATTGCTGAC 420
413 CTTCACACCACTTCTACTGAGATGAATAAATTTGATCTAGTCTCTTTATTGCTGAC 472
421 TTTTCAGACACAGCAAGAGTTAGGATTATTATCTTC 459
473 TTTTCAGACACAGCAAGAGTTAGGATTATTATCTTC 511

RESULT 4
LOCUS CQ737424 477 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 23358 from Patent WO02068579.
ACCESSION CQ737424
VERSION CQ737424.1 GI:42335713
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 23358 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
Source
1..477
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.7%; Score 453; DB 6; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.3e-115; Indels 0; Gaps 0;
Matches 453; Conservative 0; Mismatches 0;

1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCCTGATTTGGACCGGTGCTGAGAACCTC 60
1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCCTGATTTGGACCGGTGCTGAGAACCTC 60
61 CATGTGAAATAAGTTGCTCTCGACTGCTGATGGTCTCAGTTATCCAGTTGCAGAA 120
61 CATGTGAAATAAGTTGCTCTCGACTGCTGATGGTCTCAGTTATCCAGTTGCAGAA 120
121 AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
121 AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
181 CGGATACATACATATGATATGATTTATATCTTTGTCGATTTGGCATCAGGACA 240
181 CGGATACATACATGATATGATTTATATCTTTGTCGATTTGGCATCAGGACA 240
241 AGGGTAGTTCTTCAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
241 AGGGTAGTTCTTCAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
301 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360

301 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
361 CTTCACACCACTTCTACTGAGATGAATAAATTTGATCTAGTCTCTTTATTGCTGAC 420
361 CTTCACACCACTTCTACTGAGATGAATAAATTTGATCTAGTCTCTTTATTGCTGAC 420
421 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 453
421 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 453

RESULT 5
LOCUS BD109278 642 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD109278
VERSION BD109278.1 GI:23204096
KEYWORDS JP 2002010789-A/1355.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1 (bases 1 to 642)
Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
EST and encoded human protein
Patent: JP 2002010789-A 1355 15-JAN-2002;
GENSET CORP

COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/1355
PD 15-JAN-2002
PF 07-AUG-2000 JP 2002080989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 10.699998092651
CC seq VMLMLAVLIWTGA/EN
FH Key Location/Qualifiers
FT CDS 53..526
FT sig_peptide 53..103.
Location/Qualifiers
1..642
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
Source
1..642
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.7%; Score 453; DB 6; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.3e-115; Indels 0; Gaps 0;
Matches 453; Conservative 0; Mismatches 0;

1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCCTGATTTGGACCGGTGCTGAGAACCTC 60
53 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCCTGATTTGGACCGGTGCTGAGAACCTC 112
61 CATGTGAAATAAGTTGCTCTCTCGACTGCTGATGGTCTCAGTTATCCAGTTGCAGAA 120
113 CATGTGAAATAAGTTGCTCTCTCGACTGCTGATGGTCTCAGTTATCCAGTTGCAGAA 172
121 AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
173 AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 232
181 CGGATACATACATATGATATGATTTATATCTTTGTCGATTTGGCATCAGGACA 240
233 CGGATACATACATATGATATGATTTATATCTTTGTCGATTTGGCATCAGGACA 292

QY 241 AGGGTAGTTCTGAGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
293 AGGGTAGTTCTGAGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 301 ATAGATCATGACCTTCAGGAATCCATTTGGAGTGTCCACCTCTAGGAATCAGTGTGG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
353 ATAGATCATGACCTTCAGGAATCCATTTGGAGTGTCCACCTCTAGGAATCAGTGTGG 412
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAAATTTGGATCCTTAGTCTTTTATTCGTGAC 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
413 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAAATTTGGATCCTTAGTCTTTTATTCGTGAC 472
QY 421 TTTGAGACAACAGCAGAGAGTTAGGATTATTA 453
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
473 TTTGAGACAACAGCAGAGAGTTAGGATTATTA 505

RESULT 6
AR413725
LOCUS AR413725 642 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1362 from patent US 6639063.
ACCESSION AR413725
VERSION AR413725.1 GI:40168835
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 642)
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 1362 28-OCT-2003;
Genet S.A.;;
WOX;

FEATURES
source
1..642
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 98.7%; Score 453; DB 6; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
53 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 112
QY 61 CATGTGAAAATAAGTTGCTCTCTGGAAGTGTGATGCTCAGTTATCCAGTTTCAGAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 CATGTGAAAATAAGTTGCTCTCTGGAAGTGTGATGCTCAGTTATCCAGTTTCAGAA 172
QY 121 AGCAGAAATCTGTATATATTTGCGGATGAAATACATCTGGGAATGGGCTGCCCTGCAAT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 AGCAGAAATCTGTATATATTTGCGGATGAAATACATCTGGGAATGGGCTGCCCTGCAAT 232
QY 181 CGGATACATACATATGATATATATATCTTTGTCGATTTGGCATCAGGACA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
233 CGGATACATACATATGATATATATCTTTGTCGATTTGGCATCAGGACA 292
QY 241 AGGGTAGTTTCTGAGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
293 AGGGTAGTTTCTGAGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 301 ATAGATCATGACCTTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
353 ATAGATCATGACCTTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 412
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAAATTTGGATCCTTAGTCTTTTATTCGTGAC 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
413 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAAATTTGGATCCTTAGTCTTTTATTCGTGAC 472
QY 421 TTTGAGACAACAGCAGAGAGTTAGGATTATTA 453
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 473 TTTGAGACAACAGCAGAGAGTTAGGATTATTA 505
RESULT 7
AX970559
LOCUS AX970559 642 bp DNA linear PAT 15-JAN-2004
DEFINITION Sequence 1362 from Patent EP1104808.
ACCESSION AX970559
VERSION AX970559.1 GI:40978037
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: EP 1104808-A 1362 06-JUN-2001;
Genet (FR)
FEATURES
source
1..642
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
53..>526
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF15667.1"
/db_xref="GI:40978038"
/translation="MALEVLMLLAVLIWTGAENLHVKISCSLDMLMVSIVPVAESRNL
YIFADELHGMGCPRAHITVYFIVLRDCGIRTVRVSSEITLLFQELFVTPRNID
HDPOIHLECSYRSKSVMLTPVSTENEIKLDPSPFIADFTABELGLSSSPNLL"
53..103
/note="Von Heijne matrix score 10.6999998092651 seq
VIMLLAVLIWTGA/EN"

ORIGIN

Query Match 98.7%; Score 453; DB 6; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
53 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 112
QY 61 CATGTGAAAATAAGTTGCTCTCTGGAAGTGTGATGCTCAGTTATCCAGTTTCAGAA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 CATGTGAAAATAAGTTGCTCTCTGGAAGTGTGATGCTCAGTTATCCAGTTTCAGAA 172
QY 121 AGCAGAAATCTGTATATATTTGCGGATGAAATACATCTGGGAATGGGCTGCCCTGCAAT 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 AGCAGAAATCTGTATATATTTGCGGATGAAATACATCTGGGAATGGGCTGCCCTGCAAT 232
QY 181 CGGATACATACATATGATATATATATCTTTGTCGATTTGGCATCAGGACA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
233 CGGATACATACATATGATATATATCTTTGTCGATTTGGCATCAGGACA 292
QY 241 AGGGTAGTTTCTGAGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
293 AGGGTAGTTTCTGAGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 301 ATAGATCATGACCTTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
353 ATAGATCATGACCTTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 412
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAAATTTGGATCCTTAGTCTTTTATTCGTGAC 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
413 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAAATTTGGATCCTTAGTCTTTTATTCGTGAC 472
QY 421 TTTGAGACAACAGCAGAGAGTTAGGATTATTA 453
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
473 TTTGAGACAACAGCAGAGAGTTAGGATTATTA 505

```
RESULT 8
CS072271
LOCUS
DEFINITION
Sequence 19 from Patent WO2001042451.
ACCESSION
CS072271
VERSION
CS072271.1 GI:63089501
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Dumas Milne Edwards, J.B., Bougueleret, L. and Jobert, S.
TITLE
Full-length human cDNAs encoding potentially secreted proteins
JOURNAL
Patent: WO 2001042451-A 19 14-JUN-2001;
Seron Genetics Institute S.A. (FR)
FEATURES
Location/Qualifiers
1..691
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
42..518
CDS
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAI93409.1"
/db_xref="GI:63089502"
/translation="MALEVLMLLAVLIWTGAENLHVKISGLDMLVSVIPVAESRNL
YIPADELHLMGCPANRIHTYVYFIVLRDGGIRRVVSEETLLPOTELYPTPRNID
HPQETHLSCSRSKSVMLTPVSTENEIKLDPSPFIADPQTAAELGLLSSPNLL"
42..92
sig_peptide
/notes="Von Heijne matrix score 10.7019149919754 seq
VLMLLAVLIWTGA/EN"
ORIGIN
Query Match 98.7%; Score 453; DB 6; Length 691;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 60
DB 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 101
QY 61 CATGTGAAATAAGTTGCTCTCGACTGCTTGGACTGCTTGGATGCTCAGTTATCCAGTTGCAGAA 120
DB 102 CATGTGAAATAAGTTGCTCTCGACTGCTTGGACTGCTTGGATGCTCAGTTATCCAGTTGCAGAA 161
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 162 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 221
QY 181 CGGATACATACATATGTATATGATTTATATATCTTGTTCGATGCTGATTTGGCATCAGGACA 240
DB 222 CGGATACATACATATGTATATGATTTATATATCTTGTTCGATGCTGATTTGGCATCAGGACA 281
QY 241 AGGGTAGTTTCTGAGAAACTCTCCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
DB 282 AGGGTAGTTTCTGAGAAACTCTCCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 341
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 342 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 401
QY 361 CTTACACCAAGTTTCTACTCAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 420
DB 402 CTTACACCAAGTTTCTACTCAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 461
QY 421 TTTTCAGACAAACAGCAGAGTTAGGATTATTA 453
DB 462 TTTTCAGACAAACAGCAGAGTTAGGATTATTA 494
```

```
RESULT 9
AX574405
LOCUS
DEFINITION
Sequence 2 from Patent WO2060466.
ACCESSION
AX574405
VERSION
AX574405.1 GI:27551738
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Salter-Cid, L., Ebbets-Reed, D., Bour, B.A., Chicca, J., Yen-Potin, F.
and Bihain, B.
TITLE
Gsp3 polynucleotides and polypeptides and uses thereof
JOURNAL
Patent: WO 02060466-A 2 08-AUG-2002;
GENSET (FR)
FEATURES
Location/Qualifiers
1..1337
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
53..>103
CDS
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD60066.1"
/db_xref="GI:27551739"
/translation="MALEVLMLLAVLIWTGA"
104..526
sig_peptide
ORIGIN
Query Match 98.7%; Score 453; DB 6; Length 1337;
Best Local Similarity 100.0%; Pred. No. 1.2e-115;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 60
DB 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 112
QY 61 CATGTGAAATAAGTTGCTCTCGACTGCTTGGATGCTCAGTTATCCAGTTGCAGAA 120
DB 113 CATGTGAAATAAGTTGCTCTCGACTGCTTGGATGCTCAGTTATCCAGTTGCAGAA 172
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 173 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 232
QY 181 CGGATACATACATATGTATATGATTTATATATCTTGTTCGATGCTGATTTGGCATCAGGACA 240
DB 233 CGGATACATACATATGTATATGATTTATATATCTTGTTCGATGCTGATTTGGCATCAGGACA 292
QY 241 AGGGTAGTTTCTGAGAAACTCTCCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
DB 293 AGGGTAGTTTCTGAGAAACTCTCCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 353 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 361 CTTACACCAAGTTTCTACTCAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 420
DB 413 CTTACACCAAGTTTCTACTCAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 472
QY 421 TTTTCAGACAAACAGCAGAGTTAGGATTATTA 453
DB 473 TTTTCAGACAAACAGCAGAGTTAGGATTATTA 505
```

```
RESULT 10
AX600204
LOCUS
```

```
AX600204 1480 bp DNA linear PAT 14-FEB-2003
```

	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1502) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casaventi,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.D.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y., Sanchez,A., Whiting,M., Madan,A., Touchman,J.W., Green,E.D., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smalilus,D.E., Schmerch,A., Schein,J.E., Jones,S.J. and Warra,M.A.
	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
	12477932
	2 (bases 1 to 1502)
	. NIH MGC Project
	Direct Submission
	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
	NIH-MGC Project URL: http://mgc.nci.nih.gov
	Contact: MGC help desk Email: cgabbs@mail.nih.gov
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
	Web site: http://www.shgc.stanford.edu
	Contact: (Dickson, Mark) mcd@axil.stanford.edu
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov Series: IRAK Plate: 74 Row: k Column: 9 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 42476002. Location/Qualifiers 1..1502 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone=WGC:43628 IMAGE:5270384" /tissue_type="Testis" /clone_lib="NIH_MGC_97" /lab_host="DHIOB" /note="Vector: pBluescriptr"
	source 1..1502 /gene="FLJ36198" /db_xref="GeneID:219990" 52..1528 /gene="FLJ36198" /codon_start=1 /product="hypothetical protein LOC219990" /protein_id="AAH36256.1" /db_xref="GI:37537253" /db_xref="GeneID:219990" gene 1..1502 /gene="FLJ36198" /db_xref="GeneID:219990" CDS 52..1528 /gene="FLJ36198" /codon_start=1 /product="hypothetical protein LOC219990" /translation="MALEVTMLAVLIWTGAENLHVKSCLDWLMVSIVPAESRLN; YIFADHLGMGPCANIRITVEFYLVDRDCGIRTVSEETLLFTLFYYTPPRIND;

ORIGIN	HDPOEHLHCSTSRKSVMLTPVSTENEIKLDPSPFIADFTTABELGLSSPNLL"
Query Match	98.7%; Score 453; DB 8; Length 1502;
Best Local Similarity	100.0%; Pred. No. 1.2e-115;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTTGGACCGGTGCTGAGAACCTC 60
Db	52 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTTGGACCGGTGCTGAGAACCTC 111
QY	61 CATGTGAAATAAGTCTCTCTGGACTGTTGATGCTCAGTATCCAGTTGCGAGAA 120
Db	112 CATGTGAAATAAGTCTCTCTGGACTGTTGATGCTCAGTATCCAGTTGCGAGAA 171
QY	121 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 180
Db	172 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 231
QY	181 CGGATACATACATATGATGATGATTTATATATCTGTTGCTGATTTGGCATCAGGACA 240
Db	232 CGGATACATACATATGATGATGATTTATATATCTGTTGCTGATTTGGCATCAGGACA 291
QY	241 AGGTAGTCTTTCAGGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db	292 AGGTAGTCTTTCAGGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 351
QY	301 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
Db	352 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 411
QY	361 CTTACACAGTCTTCTACTGAGATGAATAAATTTGATCTGATCTGATCTTATTTGCTGAC 420
Db	412 CTTACACAGTCTTCTACTGAGATGAATAAATTTGATCTGATCTGATCTTATTTGCTGAC 471
QY	421 TTTTCAGACACAGCAGAGAGTGTAGGATTATTA 453
Db	472 TTTTCAGACACAGCAGAGAGTGTAGGATTATTA 504
RESULT 12	
BC048121	1646 bp mRNA linear PRI 28-JUL-2005
LOCUS	
DEFINITION	Homo sapiens hypothetical protein FLJ36198, mRNA (cDNA clone
ACCESSION	MGC:57403 IMAGE:5266642), complete cds.
VERSION	BC048121
KEYWORDS	BC048121.1 GI:28856186
SOURCE	MGC.
ORGANISM	Homo sapiens (human)
REFERENCE	
AUTHORS	Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Holtzman, R.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Roshlyki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
CONSRTM	Mammalian Gene Collection Program Team

TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 1646)
AUTHORS	NIH MGC Project
CONSRTM	Direct Submission
TITLE	Submitted (04-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK	Contact: MGC help desk
COMMENT	Email: gcgaps-1@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 106 Row: j Column: 14

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 42476002.

FEATURES

Location/Qualifiers

1..1646

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:57403 IMAGE:5266642"

/issue_type="Testis"

/clone_lib="NIH MGC_97"

/lab_host="DH10B"

/note="Vector: pBluescriptR"

1..1646

/gene="FLJ36198"

/db_xref="GeneID:219990"

186..662

/gene="FLJ36198"

/codon_start=1

/product="hypothetical protein LOC219990"

/protein_id="AAH48121.1"

/db_xref="GI:28856187"

/db_xref="GeneID:219990"

/translation="MALEVIMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNL
YIFADELHLMGMCPPANRHHYVYEFYLVDCGIRTVVSEETLLDFOTLYFTPRNID
HDPOEHLHCSTSRKSVMLTPVSTENEIKLDPSPFIADFTTABELGLSSPNLL"

ORIGIN

Query Match 98.7%; Score 453; DB 8; Length 1646;

Best Local Similarity 100.0%; Pred. No. 1.2e-115;

Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTTGGACCGGTGCTGAGAACCTC 60

Db 186 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTTGGACCGGTGCTGAGAACCTC 245

QY 61 CATGTGAAATAAGTCTCTCTGGACTGTTGATGCTCAGTATCCAGTTGCGAGAA 120

Db 246 CATGTGAAATAAGTCTCTCTGGACTGTTGATGCTCAGTATCCAGTTGCGAGAA 305

QY 121 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 180

Db 306 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 365

QY 181 CGGATACATACATATGATGATGATTTATATATCTGTTGCTGATTTGGCATCAGGACA 240

Db 366 CGGATACATACATATGATGATGATTTATATATCTGTTGCTGATTTGGCATCAGGACA 425

QY 241 AGGAGTATCTGAGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
 |||||
 Db 426 AGGAGTATCTGAGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 485
 |||||
 QY 301 ATAGATCATGACCCCTCAGGAAATCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
 |||||
 Db 486 ATAGATCATGACCCCTCAGGAAATCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 545
 |||||
 QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATGAAATTTGGATCCTTAGTCTCTTTTATTCTGAC 420
 |||||
 Db 546 CTTACACCAAGTTTCTACTGAGAAATGAAATGAAATTTGGATCCTTAGTCTCTTTTATTCTGAC 605
 |||||
 QY 421 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 453
 |||||
 Db 606 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 638
 |||||

RESULT 13
 BD275948 1898 bp DNA linear PAT 17-JUL-2003
 LOCUS
 DEFINITION 62 Human secreted proteins.
 ACCESSION BD275948
 VERSION BD275948.1 GI:33085716
 SOURCE JP 2002543771-A/26.
 KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1898)
 AUTHORS Birse,C.E., Mourat,P.A., Florence,K.A., Ruben,S.M.,
 Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
 Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
 TITLE 62 Human secreted proteins
 JOURNAL Patent: JP 2002543771-A 26 24-DEC-2002;
 COMMENT Human Genome Sciences Inc
 OS Homo sapiens
 PN JP 2002543771-A/26
 PD 24-DEC-2002
 PF 06-APR-2000 JP 2000611564
 PR 09-APR-1999 US 60/130991 PI
 Charles e birse,paul a mouret,kimberly a florence,steven m pi
 ruben,
 PI george a komatsoulis,jian ni,reinhard ebner,david w PI
 lafleuer,henrik s olsen,
 PI yanggu shi,daniel r soppet,craig a rosen,paul e young CC
 FH Key Location/Qualifiers
 FT SITE (1398)
 FT /notes='n equals a,t,g, or c'
 FT SITE (1428)
 FT /notes='n equals a,t,g, or c'.
 FT Location/Qualifiers
 1..1898
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source
 Query Match 98.7%; Score 453; DB 6; Length 1898;
 Best Local Similarity 100.0%; Pred. No. 1.2e-115;
 Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ORIGIN
 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
 |||||
 Db 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 74
 |||||

QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTCCAGAA 120
 |||||
 Db 75 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTCCAGAA 134
 |||||
 QY 121 AGCAGAAATCTGTATATATTTCGGATGAAATACATCTGGGAATGGGCTGCCCTCAAT 180
 |||||
 Db 135 AGCAGAAATCTGTATATATTTCGGATGAAATACATCTGGGAATGGGCTGCCCTCAAT 194
 |||||

QY 181 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 240
 |||||
 Db 195 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 254
 |||||
 QY 241 AGGAGTATCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
 |||||
 Db 255 AGGAGTATCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 314
 |||||
 QY 301 ATAGATCATGACCCCTCAGGAAATCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
 |||||
 Db 315 ATAGATCATGACCCCTCAGGAAATCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 374
 |||||
 QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATGAAATTTGGATCCTTAGTCTCTTTTATTCTGAC 420
 |||||
 Db 375 CTTACACCAAGTTTCTACTGAGAAATGAAATGAAATTTGGATCCTTAGTCTCTTTTATTCTGAC 434
 |||||
 QY 421 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 453
 |||||
 Db 435 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 467
 |||||

RESULT 14
 AX748179 1492 bp mRNA linear PAT 20-JUN-2003
 LOCUS
 DEFINITION Sequence 1704 from Patent EP1308459.
 ACCESSION AX748179
 VERSION AX748179.1 GI:32132567
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
 Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
 Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
 Masuho, Y.
 TITLE Full-length cDNA sequences
 JOURNAL Patent: EP 1308459-A 1704 07-MAY-2003;
 Helix Research Institute (JP) ; Research Association for
 Biotechnology (JP)
 FEATURES
 Location/Qualifiers
 source
 1..1492
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 98.3%; Score 451.4; DB 6; Length 1492;
 Best Local Similarity 99.8%; Pred. No. 3.3e-115;
 Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
 |||||
 Db 52 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 111
 |||||
 QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTCCAGAA 120
 |||||
 Db 112 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTCCAGAA 171
 |||||
 QY 121 AGCAGAAATCTGTATATATTTCGGATGAAATACATCTGGGAATGGGCTGCCCTCAAT 180
 |||||
 Db 172 AGCAGAAATCTGTATATATTTCGGATGAAATACATCTGGGAATGGGCTGCCCTCAAT 231
 |||||
 QY 181 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 240
 |||||
 Db 232 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 291
 |||||
 QY 241 AGGAGTATCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
 |||||
 Db 292 AGGAGTATCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 351
 |||||

```

301 ATAGATCATGACCTCAGGAATCCATTGAGGTGTTCCACCTCTAGGAATCAGTGTGG 360
|||||
352 ATAGATCATGACCTCAGGAATCCATTGAGGTGTTCCACCTCTAGGAATCAGTGTGG 411
|||||
361 CTTACACAGTTCCTACTGAGATGAATAAAATTTGGATCCTAGTCTCTTTATTTGCTGAC 420
|||||
412 CTTACACAGTTCCTACTGAGATGAATAAAATTTGGATCCTAGTCTCTTTATTTGCTGAC 471
|||||
421 TTTACAGACACAGCAGAAGATTAGGATTATTA 453
|||||
472 TTTACAGACACAGCAGAAGATTAGGATTATTA 504
|||||

RESULT 15
AK093517 1492 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ36198 fis, clone TESTI2028242, weakly similar
DEFINITION to Mus musculus EPSC26 mRNA.
ACCESSION AK093517
VERSION AK093517.1 GI:21752414
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1
Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakatsuki, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Iehibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, K., Tanase, T., Nomura, Y., Togiya, S., Konai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shichata, N., Sano, S., Mori, S., Moniama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
2
Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S.,
Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Wakatsuki, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y.,
Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished

```

```

REFERENCE
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source Location/Qualifiers
1..1492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI2028242"
/tissue_type="testis"
/clone_lib="TESTI2"
/notes="cloning vector: pME18SFL3"
52..528
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC04191.1"
/db_xref="GI:21752415"
/translation="MALEVLMLLAVLWLTGAENLHVKISGLDMLMASVIPVAESRNL
YIFADRLHGMGCPANRIHYVYEFYLYRDCGIRTRVSEETLLFTQLEYFTFPRNID
HDPQEHLECSRSRKSVMWLPVSTENEIKLDPSPFIADFTTAAELGSSSSPLL"
ORIGIN
Query Match 98.3%; Score 451.4; DB 8; Length 1492;
Best Local Similarity 99.8%; Pred. No. 3.3e-115;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGATTTGGACCGGTGCTGAGAACCTC 60
|||||
DB 52 ATGGCGTTAGAGTCTTGATGCTCTCGTCTGATTTGGACCGGTGCTGAGAACCTC 111
|||||
QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGCTCAGTTATCCAGTTGCAGAA 120
|||||
DB 112 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGCTCAGTTATCCAGTTGCAGAA 171
|||||
QY 121 AGCAGAAAATCTGTATATATTTGGGATGAATACATCTGGGATGGCTGCCCTGCAAA 180
|||||
DB 172 AGCAGAAAATCTGTATATATTTGGGATGAATACATCTGGGATGGCTGCCCTGCAAA 231
|||||
QY 181 CGGATACATACATATGATATGATGTTTATATATCTTTGTTGCGTATTTGGCATCAGGACA 240
|||||
DB 232 CGGATACATACATATGATATGATGTTTATATATCTTTGTTGCGTATTTGGCATCAGGACA 291
|||||
QY 241 AGGTTAGTTCTGAGGAAACTCTCTTTTCAACCGAGCTGTACTTACCCCAAGGAAT 300
|||||
DB 292 AGGTTAGTTCTGAGGAAACTCTCTTTTCAACCGAGCTGTACTTACCCCAAGGAAT 351
|||||
QY 301 ATAGATCATGACCTCAGGAATCCATTGAGGTGTTCCACCTCTAGGAATCAGTGTGG 360
|||||
DB 352 ATAGATCATGACCTCAGGAATCCATTGAGGTGTTCCACCTCTAGGAATCAGTGTGG 411
|||||
QY 361 CTTACACAGTTCCTACTGAGATGAATAAAATTTGGATCCTAGTCTCTTTATTTGCTGAC 420
|||||
DB 412 CTTACACAGTTCCTACTGAGATGAATAAAATTTGGATCCTAGTCTCTTTATTTGCTGAC 471
|||||
QY 421 TTTACAGACACAGCAGAAGATTAGGATTATTA 453
|||||
DB 472 TTTACAGACACAGCAGAAGATTAGGATTATTA 504
|||||
Search completed: May 9, 2006, 23:35:04
Job time : 3498.4 secs

```

This Page Blank (uspto)

Db 233 CGGATACATATATGATATGATATGATATATATCTTGTGATGATGTCATCAGGACA 292
QY 241 AGGCTAGTTCTCGAGAACTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db 293 AGGCTAGTTCTCGAGAACTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTCGAGTGTTCACCTCTAGGAATCAGTGTGG 360
Db 353 ATAGATCATGACCCCTCAGGAAATCCATTTCGAGTGTTCACCTCTAGGAATCAGTGTGG 412
QY 361 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 420
Db 413 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 472
QY 421 TTTTCAGACACAGCAGAGAGTTAGGNTATTATTA 453
Db 473 TTTTCAGACACAGCAGAGAGTTAGGNTATTATTA 505

RESULT 3

ABS78646
ID ABS78646 standard; cDNA; 1480 BP.
XX
AC ABS78646;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding, CGDD4, INCYTE 5284076CBI.
XX

Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
cell proliferative disorder; arteriosclerosis; atherosclerosis;
cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
polycythaemia vera; primary thrombocytopaenia; developmental disorder;
renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
reproductive disorder; infertility; autoimmune disorder; gout; allergy;
inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
rheumatoid arthritis.

XX
OS Homo sapiens.
XX
PN WO200272830-A2.
XX
PD 19-SEP-2002.
XX
PF 08-FEB-2002; 2002WO-US0003715.
XX
PR 09-FEB-2001; 2001US-0268111P.
PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.

XX (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR,
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Wallia NK;
XX
DR WPI: 2002-723356/78.
DR P-PSDB; ABG97353.

XX
PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX

PS Claim 5; Page 169; 181pp; English.
XX
CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino

CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX

SQ Sequence 1480 BP; 463 A; 244 C; 265 G; 508 T; 0 U; 0 Other;
Query Match 98.7%; Score 453; DB 6; Length 1480;
Best Local Similarity 100.0%; Pred. No. 3.4e-132;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGGCTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTCTGAGAACCTC 60
Db 29 ATGCGGCTAGAAAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTCTGAGAACCTC 88
QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGAGAA 120
Db 89 CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGAGAA 148
QY 121 AGCAGAAATCTGTATATATTTTCGGGATGAAATACATCTGGGATGGCTGCCCTGCAAT 180
Db 149 AGCAGAAATCTGTATATATTTTCGGGATGAAATACATCTGGGATGGCTGCCCTGCAAT 208
QY 181 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 209 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 268
QY 241 AGGCTAGTTTCTGAGGAAATCTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 300
Db 269 AGGCTAGTTTCTGAGGAAATCTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 328
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTCGAGTGTTCACCTCTAGGAATCAGTGTGG 360
Db 329 ATAGATCATGACCCCTCAGGAAATCCATTTCGAGTGTTCACCTCTAGGAATCAGTGTGG 388
QY 361 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 420
Db 389 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 448
QY 421 TTTTCAGACACAGCAGAGAGTTAGGNTATTATTA 453
Db 449 TTTTCAGACACAGCAGAGAGTTAGGNTATTATTA 481

RESULT 4
AAC69528
ID AAC69528 standard; DNA; 1898 BP.
XX
AC AAC69528;
XX

Db	435	TTTTGACAAACAGCAGAGAGTTAGGATTATTA	467
RESULT 7			
ADB91291	ID ADB91291 standard; cDNA; 1898 BP.		
XX	AC	ADB91291;	
XX	DT	04-DEC-2003 (first entry)	
XX	DE	Human secreted protein cDNA #SEQ ID 237.	
XX	KW	Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO2003004622-A2.	
XX	PD	16-JAN-2003.	
XX	PF	19-MAR-2002; 2002WO-US008124.	
XX	PR	21-MAR-2001; 2001US-0277340P.	
XX	PR	19-JUL-2001; 2001US-0306171P.	
XX	PR	13-NOV-2001; 2001US-0331287P.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Rosen CA, Ruben SM;	
XX	DE	WPI; 2003-229407/22.	
XX	PT	Nucleic acid encoding a human secreted protein is useful in diagnosing or treating diabetes or conditions related to diabetes.	
XX	PS	Claim 9; SEQ ID NO 237; 1537pp; English.	
XX	CC	The invention relates to isolated nucleic acid molecules ADB91065-ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-ADB91834. Also disclosed is a recombinant vector comprising a polynucleotide of the invention, and a recombinant host cell comprising the recombinant vector. The polypeptide of the invention is useful in identifying a binding partner by contacting the polypeptide with a binding partner, and determining whether the binding partner increases or decreases activity of the polypeptide. The polypeptide, polynucleotide, antibody or its fragment, agonist or antagonist are useful for preparing a pharmaceutical composition for diagnosing or treating diabetes or conditions related to diabetes. The present sequence is that of the human immunoglobulin Fc portion used to generate fusion proteins, increasing the stability of the fused protein as compared to the secreted protein only. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX	SQ	Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;	
Query Match 98.7%; Score 453; DB 9; Length 1898;			
Best Local Similarity 100.0%; Pred. No. 3.8e-132;			
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCGTTAGAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTCTCAGAACCTC	60
Db	15	ATGGCGTTAGAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTCTCAGAACCTC	74
QY	61	CATGTGAAATAAAGTTGCTCTCGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA	120
Db	75	CATGTGAAATAAAGTTGCTCTCGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA	134
QY	121	ACGAGAAATCTATATATTTCCGATGAAATTCATCTGGGATGGGCTGCCCTCAAT	180
Db	135	ACGAGAAATCTATATATTTCCGATGAAATTCATCTGGGATGGGCTGCCCTCAAT	194
QY	181	CGGATACATACATATGTATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA	240
Db	195	CGGATACATACATATGTATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA	254
QY	241	AGGTTAGTTTCTGAGGAAACTCTCTTTTCAACCCGAGCTGTACTTTACCCCAAGGAAT	300
Db	255	AGGTTAGTTTCTGAGGAAACTCTCTTTTCAACCCGAGCTGTACTTTACCCCAAGGAAT	314
QY	301	ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG	360
Db	315	ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG	374
QY	361	CTTACACAGTTTCTACTGAGAATGAATAAAATTTGGATCTAGTCTCTTTTATTGCTCAC	420
Db	375	CTTACACAGTTTCTACTGAGAATGAATAAAATTTGGATCTAGTCTCTTTTATTGCTCAC	434
QY	421	TTTCAGACAAACAGCAGAGAGTTAGGATTATTA	453
Db	435	TTTCAGACAAACAGCAGAGAGTTAGGATTATTA	467
RESULT 8			
ADB91291	ID ADB91291 standard; DNA; 1898 BP.		
XX	AC	ADB91291;	
XX	DT	01-JAN-2004 (first entry)	
XX	DE	Human secreted protein-related DNA - SEQ ID 356.	
XX	KW	antianaemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid; antidiabetic; immunosuppressive; dermatological; nephrotropic; antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide; fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic; haemopoietic; haematologic; anaemia; autoimmune disorder; rheumatoid arthritis; inflammation; glomerulonephritis; neurodegenerative; systemic lupus erythematosus; Grave's disease; diabetes; Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis; cancer; bacterial; viral; fungal; parasitic infection; gene therapy; human; gene; ds.	
XX	OS	Homo sapiens.	
XX	PN	WO2003038063-A2.	
XX	PD	08-MAY-2003.	
XX	PF	19-MAR-2002; 2002WO-US008277.	
XX	PR	21-MAR-2001; 2001US-0277340P.	
XX	PR	19-JUL-2001; 2001US-0306171P.	
XX	PR	13-NOV-2001; 2001US-0331287P.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Rosen CA, Ruben SM;	
XX	DE	WPI; 2003-430516/40.	
XX	DR	P-PSDB; ADB74338.	
XX	PT	New human secreted polypeptide for diagnosing, preventing or treating hematopoietic or hematologic disorders (e.g. anemia), autoimmune disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or atherosclerosis).	
XX	PS	Claim 27; SEQ ID NO 356; 2272pp; English.	
XX	CC	The invention relates to a novel human secreted polypeptide comprising a defined sequence given in the specification. The polypeptide, nucleic acid molecule, antibody, agonist or antagonist of the invention may be useful for preparing a composition for diagnosing or treating a	

Db 292 AGGGTAGTTTCTGAGAAACTCTCTTTTCAAAACGAGCTGTACTTTACCCCAAGGAAT 351
Qy 301 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 352 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 411
Qy 361 CTTACACCAAGTTTCTACTGAGAAATGAAATATAAAATGGATCCCTAGTCCCTTTTATTCTGAC 420
Db 412 CTTACACCAAGTTTCTACTGAGAAATGAAATATAAAATGGATCCCTAGTCCCTTTTATTCTGAC 471
Qy 421 TTTTCAGACACGACGAGAGTTAGGATTATTA 453
Db 472 TTTTCAGACACGACGAGAGTTAGGATTATTA 504

RESULT 10
ABZ36691
ID ABZ36691 standard; cDNA; 497 BP.
AC ABZ36691;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET coding sequence SEQ ID 564.
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2001; 2001WO-IB0000914.
XX
PR 18-APR-2001; 2001WO-IB0000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX WPI; 2003-075548/07.
XX
XX New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
PS Claim 12; Page 579; 735pp; English.

XX The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps
XX
SQ Sequence 497 BP; 127 A; 99 C; 117 G; 154 T; 0 U; 0 Other;
Query Match 97.2%; Score 446; DB 8; Length 497;
Best Local Similarity 100.0%; Pred. No. 3.5e-130;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 60
Db 52 ATGGGGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 111

Qy 61 CATGTGAAAATAAGTTGCTCTCTGGACCTGGTTGATGGTCTCAGTTATCCCAAGTTCCAGAA 120
Db 112 CATGTGAAAATAAGTTGCTCTCTGGACCTGGTTGATGGTCTCAGTTATCCCAAGTTCCAGAA 171
Qy 121 AGCAGAAATCTGTATATATTTTGGGATGAATTAATCATCTGGGAATGGGCTGCCCTGCAAAAT 180
Db 172 AGCAGAAATCTGTATATATTTTGGGATGAATTAATCATCTGGGAATGGGCTGCCCTGCAAAAT 231
Qy 181 CGGATACATACATATGATATGAGTTTATATATCTTTTCGTGATTTGGGCATCAGGACA 240
Db 232 CGGATACATACATATGATATGAGTTTATATATCTTTTCGTGATTTGGGCATCAGGACA 291
Qy 241 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 300
Db 292 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 351
Qy 301 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 352 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 411
Qy 361 CTTACACCAAGTTTCTACTGAGAAATGAAATATAAAATGGATCCCTAGTCCCTTTTATTCTGAC 420
Db 412 CTTACACCAAGTTTCTACTGAGAAATGAAATATAAAATGGATCCCTAGTCCCTTTTATTCTGAC 471
Qy 421 TTTTCAGACACGACGAGAGTTAGG 446
Db 472 TTTTCAGACACGACGAGAGTTAGG 497

RESULT 11
ABZ36690
ID ABZ36690 standard; cDNA; 500 BP.
XX
AC ABZ36690;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET coding sequence SEQ ID 563.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2001; 2001WO-IB0000914.
XX
PR 18-APR-2001; 2001WO-IB0000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX WPI; 2003-075548/07.
XX
New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
PS Claim 12; Page 578; 735pp; English.
XX
The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps
XX
SQ Sequence 497 BP; 127 A; 99 C; 117 G; 154 T; 0 U; 0 Other;
Query Match 97.2%; Score 446; DB 8; Length 497;
Best Local Similarity 100.0%; Pred. No. 3.5e-130;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 60
Db 52 ATGGGGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 111

CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps
XX
SQ Sequence 500 BP; 129 A; 100 C; 113 G; 155 T; 0 U; 3 Other;

Query Match 93.5%; Score 429; DB 8; Length 500;
Best Local Similarity 100.0%; Pred. No. 8.2e-125; Mismatches 0; Gaps 0;
Matches 429; Conservative 0; Indels 0; Gaps 0;
QY 1 ATGCGTTAGAGCTTGTGATCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 60
DB 72 ATGCGTTAGAGCTTGTGATCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 131
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTTCAGTTCATCCAGTTGCGAGAA 120
DB 132 CATGTGAAATAAGTTGCTCTCTGGACTGGTTCAGTTCATCCAGTTGCGAGAA 191
QY 121 AGCAGAAATCTGTATATATTGGGATGAATACATCTGGGATGGCTGCCCTGCAAAAT 180
DB 192 AGCAGAAATCTGTATATATTGGGATGAATACATCTGGGATGGCTGCCCTGCAAAAT 251
QY 181 CGGATACATACATATATATAGTTTATATATCTTTGTTGCTGATTTGGCATCAGGACA 240
DB 252 CGGATACATACATATATATAGTTTATATATCTTTGTTGCTGATTTGGCATCAGGACA 311
QY 241 AGGTTAGTTTCTGAGGAATCTCTCTTTTTCACACCGAGCTGATTTTACCCCAAGGAAT 300
DB 312 AGGTTAGTTTCTGAGGAATCTCTCTTTTTCACACCGAGCTGATTTTACCCCAAGGAAT 371
QY 301 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTTGG 360
DB 372 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTTGG 431
QY 361 CTTACACCATGTTTCTACTGAGAAATGAAATAAAATTCGATCTCTTTTATTTGCTGAC 420
DB 432 CTTACACCATGTTTCTACTGAGAAATGAAATAAAATTCGATCTCTTTTATTTGCTGAC 491
QY 421 TTTTCAGACA 429
DB 492 TTTTCAGACA 500

RESULT 12
AAH64748
ID AAH64748 standard; cDNA; 470 BP.
XX
AC AAH64748;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 24.
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; 58.
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX WPI; 2001-367870/38.
DR

DR P-PSDB; AAG89145.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
PS Claim 7; Page 583; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX

SQ Sequence 470 BP; 137 A; 87 C; 107 G; 139 T; 0 U; 0 Other;
Query Match 53.6%; Score 246; DB 5; Length 470;
Best Local Similarity 100.0%; Pred. No. 5.1e-67; Mismatches 0; Gaps 0;
Matches 246; Conservative 0; Indels 0; Gaps 0;
QY 1 ATGCGTTAGAGCTTGTGATCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 60
DB 137 ATGCGTTAGAGCTTGTGATCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 196
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTTCAGTTCATCCAGTTGCGAGAA 120
DB 197 CATGTGAAATAAGTTGCTCTCTGGACTGGTTCAGTTCATCCAGTTGCGAGAA 256
QY 121 AGCAGAAATCTGTATATATTGGGATGAATACATCTGGGATGGCTGCCCTGCAAAAT 180
DB 257 AGCAGAAATCTGTATATATTGGGATGAATACATCTGGGATGGCTGCCCTGCAAAAT 316
QY 181 CGGATACATACATATATATAGTTTATATATCTTTGTTGCTGATTTGGCATCAGGACA 240
DB 317 CGGATACATACATATATATAGTTTATATATCTTTGTTGCTGATTTGGCATCAGGACA 376
QY 241 AGGTTA 246
DB 377 AGGTTA 382

RESULT 13
ADJ12678
ID ADJ12678 standard; DNA; 5013 BP.
XX
AC ADJ12678;
XX
DT 20-MAY-2004 (first entry)
XX
DE DNA fragment of a BAC clone that encodes a human secreted protein Seq532.
XX
KW human; secreted; cancer; haematopoietic disease; anaemia;
KW multiple myeloma; reproductive system disorder; prostatitis;
KW inguinal hernia; musculoskeletal disease; systemic lupus erythematosus;
KW gout; cardiovascular disease; arthritis; hypernatraemia; fetal disease;
KW fetal alcohol syndrome; Down's syndrome; excretory disease;
KW urinary incontinence; renal disorder; neural; sensory disease;
KW Alzheimer's disease; meningitis; respiratory disease; emphysema;
KW occupational lung disease; endocrine disease; diabetes;
KW glomerulonephritis; digestive disease; portal hypertension;

irritable bowel syndrome; epithelial disease; scleroderma;
epidermolysis bullosa; cytostatic; antianemic; antiarthritic;
antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
antipsoriatic; antibacterial; osteopathic; dermatological; antigout;
immunomodulator; antiarrhythmic; cardiant; nootropic; antilipemic;
nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
antidiabetic; anabolic; hypertensive; vulnerary; ds.
Homo sapiens.
US2004010132-A1.
15-JAN-2004.
30-OCT-2001; 2001US-00984429.
09-OCT-1997; 97US-0061463P.
09-OCT-1997; 97US-0061527P.
09-OCT-1997; 97US-0061529P.
09-OCT-1997; 97US-0061532P.
09-OCT-1997; 97US-0061536P.
09-OCT-1997; 97US-0071498P.
08-OCT-1998; 98WO-US021142.
08-APR-1999; 99US-00288143.
01-NOV-2000; 2000US-0244591P.
(ROSE/) ROSEN C A.
(BREW/) BREWER L A.
(DUAN/) DUAN R D.
(RUBE/) RUBEN S M.
(FLOR/) FLORENCE K A.
(GORE/) GREENE J M.
(YOUN/) YOUNG P E.
(FERR/) FERRIE A M.
(YUGG/) YU G.
(FLOR/) FLORENCE C.
(EBNE/) EBNER R.
(OLSE/) OLSEN H.
Rosen CA, Brewer LA, Duan RD, Ruben SM, Florence KA, Greene JM;
Young PE, Ferrie AM, Yu G, Florence C, Ebner R, Olsen H;
WPI; 2004-090518/09.
New isolated nucleic acids and polypeptides, useful for diagnosing,
treating, preventing or ameliorating diseases or disorders e.g. cancer,
anemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's
disease.
Disclosure; SEQ ID NO 532; 286pp; English.

This invention relates to novel polynucleotides encoding human secreted
proteins. Specifically, it refers to the vectors, host cells, recombinant
and synthetic methods for producing human polynucleotides, polypeptides
and antibodies. Furthermore, it relates to screening methods to identify
agonists and antagonists that can be used to inhibit or enhance the
production and function of the secreted proteins. The present invention
describes these compositions as useful for diagnosing, treating or
preventing disorders such as cancer, haematopoietic diseases including
anaemia and multiple myeloma, reproductive system disorders including
prostatitis and inguinal hernia, musculoskeletal diseases including
systemic lupus erythematosus and gout, cardiovascular disease including
arrhythmia and hypernatraemia, mixed fetal diseases including fetal
alcohol syndrome and Down's syndrome, excretory diseases including
urinary incontinence and renal disorders, neural or sensory disease
including Alzheimer's disease and meningitis, respiratory disease
including emphysema and occupational lung disease, endocrine diseases
including diabetes and glomerulonephritis, digestive diseases including
portal hypertension and irritable bowel syndrome and connective tissue or
epithelial diseases including scleroderma and epidermolysis bullosa. As
such, there are various activities such as cytostatic, antianemic,
antiarthritic, antiasthmatic, anti-HIV, immunosuppressive,
antiinflammatory, antipsoriatic, antibacterial, osteopathic,

dermatological, antigout, immunomodulator, antiarrhythmic, cardiant,
nootropic, antilipemic, nephrotropic, uropathic, neuroprotective,
antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and
vulnerary. This polynucleotide is a DNA fragment of a BAC clone that
encodes a human secreted protein of the invention. NOTE: This sequence
does not appear in the printed specification but has been obtained in
CC electronic format from the US patent office at the following web site
CC www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.
XX
SQ Sequence 5013 BP; 1442 A; 909 C; 826 G; 1836 T; 0 U; 0 Other;
Query Match 39.9%; Score 183; DB 12; Length 5013;
Best Local Similarity 100.0%; Pred. No. 1.1e-46;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 64 GTGAAATAAGTTGCTCTCTCGACTGGTGTGATGGTCTCAGTTATCCCAAGTTGCAGAAAGC 123
DB 1 GTGAAATAAGTTGCTCTCTCGACTGGTGTGATGGTCTCAGTTATCCCAAGTTGCAGAAAGC 60
QY 124 AGAAATCTGTATATATATTTTCGGATGAATACATCTGGGAATGGGCTGCCCAATCGG 183
DB 61 AGAAATCTGTATATATATTTTCGGATGAATACATCTGGGAATGGGCTGCCCAATCGG 120
QY 184 ATACATACATATATATGAGTCTTATATATCTTGTTCGTGATTTGGCATCAGGCAAGG 243
DB 121 ATACATACATATATATGAGTCTTATATATCTTGTTCGTGATTTGGCATCAGGCAAGG 180
QY 244 GTA 246
DB 181 GTA 183
RESULT 14
ABK90052
ID ABK90052 standard; DNA; 10115 BP.
AC ABK90052;
XX
XX 05-NOV-2002 (first entry)
XX Human GSSP3 genomic sequence.
DE Human GSSP3 genomic sequence.
XX
KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
KW body mass; serum glucose regulation; body weight loss; obesity;
KW metabolic-related disorder; impaired glucose tolerance; stroke;
KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
KW hypertension; syndrome C; type I diabetes; type II diabetes;
KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
KW lipodystrophy; physical performance; exercise; dyslexia; schizophrenia;
KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200260456-A2.
XX
PD 08-AUG-2002.
XX
XX 01-FEB-2002; 2002WO-IB001333.
XX
PR 02-FEB-2001; 2001US-0266156P.
XX
PA (GEST) GENSET.
XX
XX Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
PI Bihain B;
XX
DR WPI; 2002-608487/65.
XX
PT Reducing circulating glucose levels or increasing insulin sensitivity,
PT useful for reducing body mass or preventing body weight gain, comprises
PT administering composition comprising GSSP3 polypeptide.

XX
PS Disclosure; Page 90-95; 97pp; English.
XX
CC The present invention relates to the isolation of human GSSP3
CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
CC polypeptide reduces circulating blood glucose levels, increases insulin
CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
CC polynucleotide sequences are useful in serum glucose regulation and
CC acid metabolism, body weight loss, and prevention of body weight gain.
CC Compositions comprising GSSP3 polypeptides are useful for controlling
CC blood glucose levels, for treating metabolic-related diseases or
CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
CC hyperlipidaemia, atherosclerosis, heart disease, hypertension, stroke,
CC syndrome C, type I or II diabetes, diabetes related complications,
CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
CC acanthosis nigricans, leprechaunism and lipotatrophy). The polypeptides
CC are also useful to improve physical performance during work or exercise,
CC and to treat dyslexia, attention-deficit disorder, attention-
CC deficit/hyperactivity disorder, and psychiatric disorders such as
CC schizophrenia. The present sequence encodes human GSSP3 polypeptide
XX
SQ Sequence 10115 BP; 2860 A; 1941 C;1815 G; 3475 T; 0 U; 24 Other;
S0

XX	WPI; 2001-524255/58.
XX	830 Primers useful for synthesizing full length cDNA clones and their use
XX	in genetic manipulation.
PT	
XX	Claim 2; SEQ ID NO 596; 1380pp + Sequence Listing; English.
PS	
XX	The invention relates to primers for synthesising full length cDNA
CC	clones. 830 cDNA molecules encoding a human protein have been isolated
CC	and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC	been determined. Primers for synthesising the full length cDNA are useful
CC	for clarifying the function of the protein encoded by the cDNA. The full
CC	length clones were obtained by construction of full length enriched cDNA
CC	libraries that were synthesised by the oligo-capping method. The primers
CC	enable the production of the full length cDNA easily without any special
CC	methods. The present sequence is the nucleotide sequence of the 5'-end of
CC	a cDNA provided in the invention. Note: The sequence data for this patent
CC	did not form part of the printed specification, but was obtained in CD-
CC	ROM format directly from EPO
XX	
SQ	Sequence 784 BP; 215 A; 206 C; 167 G; 190 T; 0 U; 6 Other;
	Query Match 19.7%; Score 90.2; DB 4; Length 784;
	Best Local Similarity 57.7%; Pred. No. 1e-17;
	Matches 161; Conservative 0; Mismatches 118; Indels 0; Gaps 0
Qy	8 TAGAAGTCTTGATGCTCCTCGCTCTTTGAATTGGACGGGTGCTGAGAACCTCCATGTGA 67
Db	307 TAGGACTGATGATCCTCTCACTTCGTTTTTCAGCGGTTTCAGGACAAAGTCCAATGA 366
Qy	68 AAATAAGTTGCTCTCTGCAGCTGGTTGATGGTCTCAGTTATCCCAGTTGCAGAAAAGCAGAA 127
Db	367 CTGTGCTGTGCTCCATGACTGGTTTCATGGTTCACAGTGCACCCCTTCATGCTAAACAACG 426
Qy	128 ATCTGTATATATTTGCGGATGAATACATCTGGGAATGGGCTGCCCTCCCAATCGGATAC 187
Db	427 ATGTGTGTGTACACTTTTCATGAACCTACACTTGGGCGCTGGGTTGCCCCCAAACCATGTTC 486
Qy	188 ATACATATGTATATGAGTTTATATATCTTGTTCGTGATTGGGCATCAGGACAAGGGTAG 247
Db	487 AGCACACGCTACCAGTTTACCTTACCGTGTACTGAATGTGGCATCAGGGCCAAAGCTG 546
Qy	248 TTTCCTGAGGAAACTCTCCTTTTTTCAAACCGAGCTGTACT 286
Db	547 TCTCTCAGGACATGGTTATCTACAGCACTGAGATACACT 585

Search completed: May 9, 2006, 14:08:21
Job time : 592.3 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:55:51 ; Search time 146.7 Seconds
(without alignments)
5561.696 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511

Perfect score: 459

Sequence: 1 atggcgttagaagctctgtat.....agtaggattataatcttc 459

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
- 7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	459	100.0	512	3	US-09-621-976-43
2	453	98.7	642	3	US-09-621-976-1362
3	451.4	98.3	1492	3	US-10-104-047-1704
4	341.6	74.4	560	3	US-09-621-976-2587
5	272	59.3	372	3	US-09-621-976-15700
6	246	53.6	469	3	US-09-621-976-1363
7	90.2	19.7	1129	3	US-09-949-016-5361
8	90.2	19.7	96590	3	US-09-949-016-17103
9	45.4	9.9	1098	3	US-09-484-9708-55
10	37.2	8.1	84252	3	US-09-949-016-17315
11	36.6	8.0	1125	3	US-09-248-796A-7831
12	35.4	7.7	53451	3	US-09-949-016-15859
13	35	7.6	49848	3	US-09-949-016-15675
14	34.8	7.6	5006	2	US-08-485-588-2
15	34.8	7.6	5006	2	US-08-484-565-2
16	34.8	7.6	5006	2	US-08-480-751-2
17	34.8	7.6	5006	2	US-08-943-986-2
18	34.8	7.6	5006	2	US-08-353-784-2
19	34.8	7.6	5006	3	US-08-484-7198-2
20	34.8	7.6	5006	3	US-08-546-998-1
21	34.8	7.6	5006	3	US-08-484-159-2
22	34	7.4	186734	3	US-09-949-016-14870
23	34	7.4	193689	3	US-09-949-016-12350
24	34	7.4	193689	3	US-09-949-016-13088

ALIGNMENTS

RESULT 1

US-09-621-976-43
; Sequence 43, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 43
; TYPE: DNA
; LENGTH: 512
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..511
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.6999998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-621-976-43

Query Match 100.0%; Score 459; DB 3; Length 512;

Best Local Similarity 100.0%; Pred. No. 1.5e-134; Mismatches 0; Indels 0; Gaps 0;

Matches 459; Conservative 0;	
QY	1 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
DB	53 ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 112
QY	61 CATGTGAAAAAAGTGTCTCTCGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 120
DB	113 CATGTGAAAAAAGTGTCTCTCGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 172
QY	121 AGCAGAAATCTGTATATATTTTGGCGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
DB	173 AGCAGAAATCTGTATATATTTTGGCGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 232
QY	181 CGGATACATACATATGTATATGAGTTTATATATCTTTGTTGATTGTTGGCATCAGGACA 240
DB	233 CGGATACATACATATGTATATGAGTTTATATATCTTTGTTGATTGTTGGCATCAGGACA 292

QY	361	CTTTACACAGCTTTCTACTAGAGAAATAAAATGGATCCTTACGTCCTTTTATCTGTGAC	420
Db	413	CTTTACACAGCTTTCTACTAGAGAAATAAAATGGATCCTTACGTCCTTTTATCTGTGAC	472
QY	421	TTTTCAGACAAACAGCAGAGAAAGATTAGGATTATTTA	453
Db	473	TTTTCAGACAAACAGCAGAGAAAGATTAGGATTATTTA	505

```

, CURRENT APPLICATION NUMBER: US/10/104,047
, CURRENT FILING DATE: 2002-03-25
, PRIOR APPLICATION NUMBER:
, PRIOR FILING DATE:
, NUMBER OF SEQ ID NOS: 4096
, SOFTWARE: PatentIn Ver. 2.1
, SEQ ID NO 1704
, LENGTH: 1492
, TYPE: DNA
, ORGANISM: Homo sapiens
US-10-104-047-1704

```

Accession	Sequence	Length
Qy	CGGATACATACATATGTTATATGAGTTTATATATCTTTGTCGTGATGTTGCGATCAGGACA	240
Db	CGGATACATACATATGTTATGAGTTTATATCTTTGTCGTGATGTTGCGATCAGGACA	291

Qy	241	AGGGTAGTTTCTGAGGAAAC	TCTCTCTTTTTC	CAACCGAGCTGTGTA	CTTTTACCCCAAGGAAT	300
Db	292	AGGGTAGTTTCTGAGGAAAC	TCTCTCTTTTTC	CAACCGAGCTGTGTA	CTTTTACCCCAAGGAAT	351
Qy	301	ATAGATCATGACCCCTCAGGAAT	CCATTTGGAGTGTCC	ACCTCTAGGAAATCAGTGTGG	360	
Db	352	ATAGATCATGACCCCTCAGGAAT	CCATTTGGAGTGTCC	ACCTCTAGGAAATCAGTGTGG	411	
Qy	361	CTTACACCAGTTTCTACTGAGAA	TGAAATATAA	TTGGATCCTAGTCC	TTTTATTGCTGAC	420
Db	412	CTTACACCAGTTTCTACTGAGAA	TGAAATATAA	TTGGATCCTAGTCC	TTTTATTGCTGAC	471
Qy	421	TTTCAGACAA	CAGCAGAGAGT	TTAGGATTATTTA	453	
Db	472	TTTCAGACAA	CAGCAGAGAGT	TTAGGATTATTTA	504	

RESULT 4
US-09-621-976-2587
; Sequence 2587, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2587
LENGTH: 560
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: CDS
LOCATION: 372..539
US-09-621-976-2587

Query Match 74.4%; Score 341.6; DB 3; Length 560;
Best Local Similarity 93.1%; Pred. No. 1.5e-97;
Matches 434; Conservative 10; Mismatches 9; Indels 13; Gaps 8;

QY 1 ATGGCGTTAGAACTCTGTATAT-ATTTCGGATGAATTACATC---TGGGAATGGGCTGCC 172
DB 53 ATGGCGTTAGAACTCTGTATAT-ATTTCGGATGAATTACATC---TGGGAATGGGCTGCC 232
QY 61 CATGTGAAAATAAGTTGCTCTCTGG---ACTGGTTGATGGTCTCAGTTATCCAGTTGCA 117
DB 113 CATGTGAAAATAAGTTGCTCTCTGGGACTTGGTTGATGGTCTCAGTTATCCAGTTGCA 172
QY 118 GAAAGCAGAAATCTGTATAT-ATTTCGGATGAATTACATC---TGGGAATGGGCTGCC 172
DB 173 GAAAGCAGAAATCTGTATAT-ATTTCGGATGAATTACATC---TGGGAATGGGCTGCC 232
QY 173 CTCGAAATCGG-ATACATACATATGATATG-AGTTTATATATCTTTGTCGATGTTGG 230
DB 233 CTCGAAATCGGATACATACATATGATATGATGAAAGTTTATATATCTTTGTCGATGTTGG 292
QY 231 CATCAGSACAAGG-TAGTTTCTGAGAACTCTCTCTTTTCAACCGAGCTGTACTTTA 289
DB 293 CATCAGSACAAGG-TAGTTTCTGAGAACTCTCTCTTTTCAACCGAGCTGTACTTTA 352
QY 290 CCCC-AAGGAATATAGATCATGACCCCTCAGGAATCCATTT-GGAGTGTTCACCTCTAG 347
DB 353 CCCC-AAGGAATATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAG 412
QY 348 GAAATCAGTGTGCTTACACCACTTTCTACTGAGATGAATAAATAAATGGATCCTAGTCC 407
DB 413 GAAATCAGTGTGCTTACACCACTTTCTACTGAGATGAATAAATAAATGGATCCTAGTCC 472
QY 408 TTTTATTGCTGACTTTTCAGACACAGCAGAGAGTTAGGATTATTA 453
DB 473 TTTTATTGCTGACTTTTCAGACACAGCAGAGAGTTAGGATTATTA 518

RESULT 5
US-09-621-976-15700
Sequence 15700, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 15700
LENGTH: 372
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc_feature
LOCATION: 351
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15700

Query Match 59.3%; Score 272; DB 3; Length 372;
Best Local Similarity 98.7%; Pred. No. 1.1e-75;
Matches 294; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

QY 1 ATGGCGTTAGAACTCTGTATATCTCTCGCTGCTCTGTGATTTGGACCGGTGCTGAGAACCTC 60
DB 76 ATGGCGTTAGAACTCTGTATATCTCTCGCTGCTCTGTGATTTGGACCGGTGCTGAGAACCTC 134
QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAAA 120
DB 135 CATGTGAAAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAAA 194
QY 121 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 195 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 254
QY 181 CGGATACATACATATCTATATATGATTTTATATATCTTTGTTGTTGTCATCAGGACA 240
DB 255 CGGATACATACATATCTATATATGATTTTATATATCTTTGTTGTTGTCATCAGGACA 314
QY 241 AGGCTAGTTTCTGAGGAAACTCTCCTTTTCAAACC-GAGCTGTACTTTTACCCCAAGG 297
DB 315 AGGCTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTTACCCCAAGG 372

RESULT 6
US-09-621-976-1363
Sequence 1363, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 1363
LENGTH: 469
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 137..454
NAME/KEY: sig_peptide
LOCATION: 137..187
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 10.699998092651
OTHER INFORMATION: seq VLMLLAVLWTGA/EN
US-09-621-976-1363

Query Match 53.6%; Score 246; DB 3; Length 469;
Best Local Similarity 100.0%; Pred. No. 1.9e-67;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAACTCTGTATATCTCTCGCTGCTCTGTGATTTGGACCGGTGCTGAGAACCTC 60
DB 137 ATGGCGTTAGAACTCTGTATATCTCTCGCTGCTCTGTGATTTGGACCGGTGCTGAGAACCTC 196
QY 61 CATGTGAAAATAAGTTGCTCTCTGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAAA 120
DB 197 CATGTGAAAATAAGTTGCTCTCTGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAAA 256
QY 121 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 257 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 316

```
QY 181 CGGATACATACATATGATGTTTATATATCTTGTTCGTGATTGGCATCAGGACA 240
|||||
Db 317 CGGATACATACATATGATGTTTATATATCTTGTTCGTGATTGGCATCAGGACA 376
|||||
QY 241 AGGGTA 246
|||||
Db 377 AGGGTA 382
|||||

RESULT 7
US-09-949-016-5361
; Sequence 5361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5361
; LENGTH: 1129
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5361

Query Match 19.7%; Score 90.2; DB 3; Length 1129;
Best Local Similarity 57.7%; Pred. No. 3.4e-18;
Matches 161; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 8 TAGAAGTCTTGATGCTCTCGCTGCTTGATTTGGACCGGTGCTGAGAACCTCCATGTGA 67
|||||
Db 309 TAGGACTGATGATCTCTCTCACCTTGTGCGTTTTCAGCCGGTTCAGGACAAAGTCCAATGA 368
|||||
QY 68 AAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAA 127
|||||
Db 369 CTGTGCTGTGCTCCATAGACTGGTTTCATGGTCACAGTGCACCCCTTCATGCTTAACAACG 428
|||||
QY 128 ATCTGTATATTTCCGGATGAATTACATCTGGGAATGGCTGCCCTGCAAAATCGGATAC 187
|||||
Db 429 ATGTGTGTGTACACTTTCATGAACCTACACTTGGGCTGGGTTGCCCCCAAAACCATGTTC 488
|||||
QY 188 ATACATATGATGATGTTTATATCTTGTTCGTGATTGGCATCAGGACAAGGTTAG 247
|||||
Db 489 AGCCACACGCTTACAGTTTCACTACCGTGTACTGAAATGTGGCATCAGGCGCAAGCTG 548
|||||
QY 248 TTTCTGAGAACTCTCTCTTTTCAAAACCGAGCTGTACT 286
|||||
Db 549 TCTCTCAGGACATGGTTATCTACAGCACTGAGATACACT 587
|||||

RESULT 8
US-09-949-016-17103
; Sequence 17103, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17103
; LENGTH: 9690
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17103

Query Match 19.7%; Score 90.2; DB 3; Length 9690;
Best Local Similarity 57.7%; Pred. No. 2.3e-17;
Matches 161; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 8 TAGAAGTCTTGATGCTCTCGCTGCTTGATTTGGACCGGTGCTGAGAACCTCCATGTGA 67
|||||
Db 93869 TAGGACTGATGATCTCTCTCACCTTGTGCGTTTTCAGCCGGTTCAGGACAAAGTCCAATGA 93928
|||||
QY 68 AAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAA 127
|||||
Db 93929 CTGTGCTGTGCTCCATAGACTGGTTTCATGGTCACAGTGCACCCCTTCATGCTTAACAACG 93988
|||||
QY 128 ATCTGTATATTTCCGGATGAATTACATCTGGGAATGGCTGCCCTGCAAAATCGGATAC 187
|||||
Db 93989 ATGTGTGTGTACACTTTCATGAACCTACACTTGGGCTGGGTTGCCCCCAAAACCATGTTC 94048
|||||
QY 188 ATACATATGATGATGTTTATATCTTGTTCGTGATTGGCATCAGGACAAGGTTAG 247
|||||
Db 94049 AGCCACACGCTTACAGTTTCACTACCGTGTACTGAAATGTGGCATCAGGCGCAAGCTG 94108
|||||
QY 248 TTTCTGAGAACTCTCTCTTTTCAAAACCGAGCTGTACT 286
|||||
Db 94109 TCTCTCAGGACATGGTTATCTACAGCACTGAGATACACT 94147
|||||

RESULT 9
US-09-484-970B-55
; Sequence 55, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 55
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 200512.1
; NAME/KEY: unsure
; LOCATION: 927, 961, 987, 1048
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-55

Query Match 9.9%; Score 45.4; DB 3; Length 1098;
Best Local Similarity 54.8%; Pred. No. 0.00044;
Matches 153; Conservative 0; Mismatches 121; Indels 5; Gaps 3;

QY 8 TAGAAGTCTTGATGCTCTCGCTGCTTGATTTGGACCGGTGCTGAGAACCTCCATGTGA 67
|||||
Db 279 TAGGACTGATGATCTCTCTCACCTTGTGCGTTTTCAGCCGGTTCAGGACAAAGTCCAATGA 338
|||||
QY 68 AAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAA 127
|||||
```


Db 339 CTGTCTGTGCTCCATAGACTGGTTTCATGGTCCACAGTGCACCCCTTCATGCTAAA-CAACG 397
QY 128 ATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGGCAATCGATAC 187
Db 398 ATGTGTGTACACTTTTCATGACT--ACAATGGGCTGGGTGGCCCCCAACCATGTTTC 454
QY 188 ATACATATGTATATGAGTTTATATATCTTCTGATTTGTGGCATCAGGACAAGGGTAG 247
Db 455 AGCCACAGCCTACCACTTCA-CTACCGTGTACTGAATGTGGCAITCAGGGCCAAAGCTG 513
QY 248 TTTCTGAGGAACTCTCCTTTTCAACCCGAGCTGTA 286
Db 514 TCTCTCAGGACATGGTTTATCTACAGCACTGAGATACACT 552

RESULT 10
US-09-949-016-17315
; Sequence 17315, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17315
; LENGTH: 84252
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17315

Query Match 8.1%; Score 37.2; DB 3; Length 84252;
Best Local Similarity 53.1%; Pred. No. 1.1;
Matches 102; Conservative 0; Mismatches 88; Indels 2; Gaps 1;

QY 174 TCGAAATCGATACATACATATATATGATGATTTATATCTTTGTTGTTGATTTGGCAT 233
Db 58068 TGTATTTTAAATATAATATTAATATATAATAAAGTAGCACCCCTTTTATGGCAT 58127
QY 234 CAGGACACAGGAGTGTCTGAGGAACTCTCTTTTCAACCGAGCTGTACTTTACCCC 293
Db 58128 ATGAAGAGAATAGCTTCTTAATACTATTTTGTAAATTAATATTCACCTTTCTCAT 58187
QY 294 AAGGAATATAGATCATG--ACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGAAA 351
Db 58188 TTTTATTTGATCTGTTATCTCTAGTAAAGAGTAACGTATCTCTAGAAATAGAAA 58247
QY 352 TCAGTGTGGCTT 363
Db 58248 TTTGAGTTGCTT 58259

RESULT 11
US-09-248-796A-7831
; Sequence 7831, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7831
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7831

Query Match 8.0%; Score 36.6; DB 3; Length 1125;
Best Local Similarity 51.5%; Pred. No. 0.26;
Matches 84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 274 ACGAGCTGTACTTTTACCCCAAGGAATATAGATCATGACCCCTCAGGAATCCATTTGGAG 333
Db 694 AACTGGATGTCAATCAATCCCTCAGTTTAAACGATAGTGAATAATGAAATATATATAGAT 753
QY 334 TGTTCACCTCTAGGAATCAGTGTGGCTTACACCAAGTTTCTACTGAGATGAATAAAA 393
Db 754 TGTATCTCTGGTGTATATCAACAGTGTCAACCAAGCACTTCTGACGAAGAAATTTGAT 813
QY 394 TTGGATCTTAGTCCCTTTTATTTGCTGACTTTTCAGACAAACAGCAG 436
Db 814 TTGGCAGATATTGAGCAGCATATTGGAATTACAGCTAGAAAAG 856

RESULT 12
US-09-949-016-15859
; Sequence 15859, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15859
; LENGTH: 53451
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15859

Query Match 7.7%; Score 35.4; DB 3; Length 53451;
Best Local Similarity 61.3%; Pred. No. 3.3; Mismatches 36; Indels 0; Gaps 0;

QY 127 AATCTGTATATATTTCCGATGAATTACATCTTGGGAATGGCTGCCCTCAATCCGATA 186
Db 28316 AATATGTATGTTTTTAAATTGTAAGGATATCTTTTATAAGTCTACCATCAACTCAATA 28375
QY 187 CATACATATGATATGAGTTTATATATCTTGT 219
Db 28376 TATTCACATGATATGTTTGTGTTTGTGTTTGT 28408

RESULT 13
US-09-949-016-15675/c
; Sequence 15675, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15675
LENGTH: 49848
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(49848)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15675

Query Match 7.6%; Score 35; DB 3; Length 49848;
Best Local Similarity 52.4%; Pred. NO. 4.3;
Matches 77; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 66 GAAATAAGTCTCTCTCGACTGGTTGATGCTCAGTTATCCCGAGTTGCAGAAAGCAG 125
|||||
Db 11593 GAATATATGTTATGTCATATATATGTCATGATACATATATACCTATGTCATGAT 11534
|||||

QY 126 AATCTGTATATATTTGGCGATGAATTACATCTGGGAATGGCTGCCCTCGCAATCGGAT 185
|||||
Db 11533 ACATATACATATGTTATGATATATATATATATACGTATATGTTATGATATATATATAT 11474
|||||

QY 186 ACATACATATGTTATGATTTATATA 212
|||||
Db 11473 ATACATATATGTTATGATACATATATA 11447
|||||

RESULT 14
US-08-485-588-2
Sequence 2, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994

Query Match 7.6%; Score 34.8; DB 2; Length 5006;
Best Local Similarity 52.0%; Pred. No. 1.8;
Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 283 TACTTTACCCCAAGAAATATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCAC 342
|||||
Db 3780 TCCTTTCTCTGAGGAGAGAGGATATAGACATCAATGCCCGCAATTTAGTCACAC 3839
|||||

QY 343 TCTAGGAATCAGTGTGGCTTACACCACTTCTCTAGAGATGAATAAATTTGATCCT 402
|||||
Db 3840 CATCTTAAATGACAGTGAATTGACCCATGTTCCCTTTAAATTTAAAAAAGAGAGCCT 3899
|||||

QY 403 AGTCCTTTTATTGCTGACTTTTCAGACAACA 432
|||||
Db 3900 TGTGTTCTGTTGTCATTTGTCAAGCA 3929
|||||

RESULT 15
US-08-484-565-2
Sequence 2, Application US/08484565
Patent No. 5763569
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:56:11 ; Search time 1220.4 Seconds
(without alignments)
3110.164 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511
Perfect score: 459
Sequence: 1 atggcgtagaagctctgat.....agttagattattattcttc 459

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext, 1.0

Searched: 9793542 seqs, 413469005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/us08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/us09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/us09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/us10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/us10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/us10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/us10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/us10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/us11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453	98.7	691	3	US-09-731-872-19 Sequence 19, Appl
2	453	98.7	691	3	US-09-876-997-19 Sequence 19, Appl
3	453	98.7	691	9	US-10-643-836-19 Sequence 19, Appl
4	453	98.7	1337	8	US-10-467-046-2 Sequence 2, Appl
5	453	98.7	1480	7	US-10-467-535-16 Sequence 16, Appl
6	453	98.7	1898	5	US-10-050-704-27 Sequence 27, Appl
7	453	98.7	1898	9	US-10-798-512-27 Sequence 27, Appl
8	453	98.7	1898	9	US-10-472-533-175 Sequence 175, Appl
9	451.4	98.3	1492	6	US-10-104-047-1704 Sequence 1704, Appl
10	246	53.6	470	3	US-09-731-872-24 Sequence 24, Appl
11	246	53.6	470	3	US-09-876-997-24 Sequence 24, Appl
12	246	53.6	470	9	US-10-643-836-24 Sequence 24, Appl
13	183	39.9	5013	3	US-09-984-429-532 Sequence 532, Appl
14	183	39.9	10115	8	US-10-467-046-1 Sequence 1, Appl
15	90.2	19.7	838	7	US-10-287-971-39 Sequence 39, Appl
16	90.2	19.7	1148	6	US-10-411-224-36 Sequence 36, Appl
17	90.2	19.7	1148	6	US-10-047-021-36 Sequence 36, Appl
18	90.2	19.7	1148	9	US-10-970-429-36 Sequence 36, Appl
19	90.2	19.7	1153	3	US-09-745-763-202 Sequence 202, Appl
20	90.2	19.7	1187	3	US-09-996-952-4 Sequence 4, Appl
21	81.4	17.7	697	7	US-10-287-971-41 Sequence 41, Appl
22	81	17.6	386	5	US-10-040-733-196 Sequence 196, Appl
23	72.4	15.8	589	8	US-10-333-872A-44 Sequence 44, Appl

ALIGNMENTS

RESULT 1

US-09-731-872-19
; Sequence 19, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouquelaret, Lydie
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-731-872-19

Query Match 98.7%; Score 453; DB 3; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.8e-133;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTTC	60
Db	42	ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTTC	101
QY	61	CATGTGAAAAAATGAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA	120
Db	102	CATGTGAAAAAATGAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA	161
QY	121	AGCAGAAATCTGTATATATTTTCGGAATGAATTCATCTGGGAATGGCTGCCCTGCAAAAT	180

Db 162 AGCAGAAATCTGTATATATTTCCGGATGAATTAACATCTGGGAATGGGCTGCCCTGCAAAAT 221
QY 181 CGGATACATACATATGATATATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 240
Db 222 CGGATACATACATATGATATATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 281
QY 241 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db 282 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 341
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
Db 342 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 401
QY 361 CTTACACAGTTTCTACTGAGAAATGAAATAAAATTTGGATCCTTAGTCCCTTTATTTGCTGAC 420
Db 402 CTTACACAGTTTCTACTGAGAAATGAAATAAAATTTGGATCCTTAGTCCCTTTATTTGCTGAC 461
QY 421 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 453
Db 462 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 494

RESULT 2

US-09-876-997-19
; Sequence 19, Application US/09876997
; Publication No. US20030152921A1

GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-09-876-997-19

Query Match 98.7%; Score 453; DB 3; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.8e-133;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGTCCTTGATTTGGACCGGTGCTCAGAACCTC 60
Db 42 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGTCCTTGATTTGGACCGGTGCTCAGAACCTC 101
QY 61 CATGTGAAATAAAGTTGCTCTCTGGACTGGTTGATGCTCAGTTATCCAGTTGCAGAA 120
Db 102 CATGTGAAATAAAGTTGCTCTCTGGACTGGTTGATGCTCAGTTATCCAGTTGCAGAA 161
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATTAACATCTGGGAATGGGCTGCCCTGCAAAAT 180
Db 162 AGCAGAAATCTGTATATATTTCCGGATGAATTAACATCTGGGAATGGGCTGCCCTGCAAAAT 221

QY 181 CGGATACATACATATGATATATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 240
Db 222 CGGATACATACATATGATATATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 281
QY 241 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db 282 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 341
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
Db 342 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 401
QY 361 CTTACACAGTTTCTACTGAGAAATGAAATAAAATTTGGATCCTTAGTCCCTTTATTTGCTGAC 420
Db 402 CTTACACAGTTTCTACTGAGAAATGAAATAAAATTTGGATCCTTAGTCCCTTTATTTGCTGAC 461
QY 421 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 453
Db 462 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 494

RESULT 3

US-10-643-836-19
; Sequence 19, Application US/10643836
; Publication No. US20050096458A1

GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN
US-10-643-836-19

Query Match 98.7%; Score 453; DB 9; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.8e-133;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGTCCTTGATTTGGACCGGTGCTCAGAACCTC 60
Db 42 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGTCCTTGATTTGGACCGGTGCTCAGAACCTC 101
QY 61 CATGTGAAATAAAGTTGCTCTCTGACTGGTTGATGCTCAGTTATCCAGTTGCAGAA 120
Db 102 CATGTGAAATAAAGTTGCTCTCTGACTGGTTGATGCTCAGTTATCCAGTTGCAGAA 161
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATTAACATCTGGGAATGGGCTGCCCTGCAAAAT 180
Db 162 AGCAGAAATCTGTATATATTTCCGGATGAATTAACATCTGGGAATGGGCTGCCCTGCAAAAT 221
QY 181 CGGATACATACATATGATATATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 240

[illegible]

```

RESULT 4
US-10-467-046-2
: Sequence 2, Application US/10467046
: Publication No. US20040235709A1
: GENERAL INFORMATION:
: APPLICANT: Salter-Cid, Luisa
: APPLICANT: Ebbets-Reed, Dana
: APPLICANT: Chicca, Barbara A.
: APPLICANT: Chicca, John
: APPLICANT: Yen-Potin, Frances
: APPLICANT: Bihain, Bernard
: TITLE OF INVENTION: GSP3 Polynucleotides and Polypeptides and Uses Thereof
: FILE REFERENCE: G-102US03REG
: CURRENT APPLICATION NUMBER: US/10/467,046
: CURRENT FILING DATE: 2003-08-01
: PRIOR APPLICATION NUMBER: US 60/266,156
: PRIOR FILING DATE: 2001-02-02
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 1337
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (53)..(103)
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: (104)..(526)
: US-10-467-046-2

```

Query Match	98.7%;	Score 453;	DB 8;	Length 1337;
Best Local Similarity	100.0%;	Pred. No. 3.9e-133;		
Matches 453;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGCGGTTAGAAAGCTTTGATGCTCCTCGCTGTCCTTGATTTGGACCGGTCGTGAGAACCTC	60	
Db	53	ATGGCGTTAGAGGCTTTGATGCTCCTCGCTGTCCTTGATTTGGACCGGTCGTGAGAACCTC	112	
QY	61	CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGCTCTCAGTTATCCCAAGTTGCAGAA	120	
Db	113	CATGTGAAAATAAGTTGCTCTCTGGA CTGGTTGATGCTCTCAGTTATCCCAAGTTGCAGAA	172	
QY	121	AGCAGAAATCTGTATATATATTGGCGATGAAATTACATCTGGGAATGGGCTGCCCTGC	180	
Db	173	AGCAGAAATCTGTATATATATTGGCGATGAAATTACATCTGGGAATGGGCTGCCCTGC	232	
QY	181	CGGATACATACATATGTATATAGCTTTATATATATCTCTGTTCTGTGATCTGGCATCAGGACA	240	
Db	233	CGGATACATACATATGTATATAGCTTTATATATATCTCTGTTCTGTGATCTGGCATCAGGACA	292	
QY	241	AGGGTAGTTTCTGAGGAAACTCTCCTTTTTTCAACCCGAGCTGTACTTTACCCCAAGGAAT	300	

Db	293	AGGTAGTTCCTGAGGAACCTCTCCTCTTTTCAAAACCGAGCTGTACTTTTACCCCAAGGAAT	352
Qy	301	ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTCCACCTCTAGAGAAATCAGTGTGG	360
Db	353	ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTCCACCTCTAGGAATCAGTGTGG	412
Qy	361	CTTACACCAGTTTCTACTGGAATGAAATAAAATGGATCCTAGTCTCTTTTATTCTGTAC	420
Db	413	CTTACACCAGTTTCTACTGGAATGAAATAAAATGGATCCTAGTCTCTTTTATTCTGTAC	472
Qy	421	TTTCAGACAACAGCAGAGAGCTTAGGATTATTA	453
Db	473	TTTCAGACAACAGCAGAGAGCTTAGGATTATTA	505

RESULT 5

```

RES001 3
US-10-467-535-16
: Sequence 16, Application US/10467535
: Publication No. US20040146970A1
: GENERAL INFORMATION:
: APPLICANT: YUE, Henry; YAO, Monique G.
: APPLICANT: ISON, Craig H.; LU, Yan
: APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
: APPLICANT: BAUGHN, Mariah R.; DING, Li
: APPLICANT: XU Yuming; GIETZEN, Kimberly J.
: APPLICANT: TANG, Tom Y.; LAL, Preeti G.
: APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
: APPLICANT: LU, Dyoung Aina M.; RICHARDSON, Thomas W.
: APPLICANT: TRAN, Uyen K.; KHARE, Reena
: APPLICANT: CHAWLA, Narinder K.
: TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
: FILE REFERENCE: PF-0903 USN
: CURRENT APPLICATION NUMBER: US/10/467,535
: CURRENT FILING DATE: 2003-08-08
: PRIOR APPLICATION NUMBER: PCT/US02/03715
: PRIOR FILING DATE: 2002-08-02
: PRIOR APPLICATION NUMBER: US 60/269,111
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: US 60/271,175
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: US 60/274,503
: PRIOR FILING DATE: 2001-03-08
: PRIOR APPLICATION NUMBER: US 60/274,552
: PRIOR FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PERL Program
: SEQ ID NO 16
: LENGTH: 1480
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No: 5284076CB1
US-10-467-535-16

```

Query Match	98.7%;	Score 453;	DB 7;	Length 1480;
Best Local Similarity	100.0%;	Prod. No. 4.1e-133;		
Matches 453;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCGGTTAGAAGCTTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCTGAGAACCTC	60	
Db	29	ATGCGGTTAGAAGCTTTGATGCTCCTCGCTGCTTGATTTGGACCGGTGCTGAGAACCTC	88	
Qy	61	CATGTGAAATAAGTTGCTCTCTGGAGCTGGTTCGATGGTCTCAGTTATCCCAAGTTGCGAAA	120	
Db	89	CATGTGAAATAAGTTGCTCTCTGGAGCTGGTTCGATGGTCTCAGTTATCCCAAGTTGCGAAA	148	
Qy	121	AGCAGAAATCTGTATATATTTGGCGAGTAATTCATCTGGGAATGGGCTGCCCTGCAAAAT	180	
Db	149	AGCAGAAATCTGTATATATTTGGCGAGTAATTCATCTGGGAATGGGCTGCCCTGCAAAAT	208	
Qy	181	CGGATACATACATATGTATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA	240	

Db 209 CGGATACATACATATGTATATAGTTTATATATCTTGTTCTGATGTTGTGGCATCAGGACA 268
QY 241 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTACCCCAAGGAAT 300
Db 269 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTACCCCAAGGAAT 328
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCCTAGGAAATCAGTGTGG 360
Db 329 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCCTAGGAAATCAGTGTGG 388
QY 361 CTTACACAGTTTCTACTGAGATGAATAAAATTTGGATCCTAGTCCCTTTATTGCTGAC 420
Db 389 CTTACACAGTTTCTACTGAGATGAATAAAATTTGGATCCTAGTCCCTTTATTGCTGAC 448
QY 421 TTTACAGCAACAGCAGAGAGTTAGGATTATTA 453
Db 449 TTTACAGCAACAGCAGAGAGTTAGGATTATTA 481

RESULT 6
US-10-050-704-27
; Sequence 27, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-050-704-27

Query Match 98.7%; Score 453; DB 5; Length 1898;
Best Local Similarity 100.0%; Pred. No. 4.7e-133;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 60
Db 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 74
QY 61 CATGTGAAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 75 CATGTGAAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 134
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 180
Db 135 AGCAGAAATCTGTATATATTTCCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 194
QY 181 CGGATACATACATATGATGATTTATATCTTGTTCGTGATTTGGCATCAGGACA 240
Db 195 CGGATACATACATATGATGATTTATATCTTGTTCGTGATTTGGCATCAGGACA 254
QY 241 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300

Db 255 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 314
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCCTAGGAAATCAGTGTGG 360
Db 315 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCCTAGGAAATCAGTGTGG 374
QY 361 CTTACACAGTTTCTACTGAGATGAATAAAATTTGGATCCTAGTCCCTTTATTGCTGAC 420
Db 375 CTTACACAGTTTCTACTGAGATGAATAAAATTTGGATCCTAGTCCCTTTATTGCTGAC 434
QY 421 TTTACAGCAACAGCAGAGAGTTAGGATTATTA 453
Db 435 TTTACAGCAACAGCAGAGAGTTAGGATTATTA 467

RESULT 7
US-10-798-512-27
; Sequence 27, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; PRIOR FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-798-512-27

Query Match 98.7%; Score 453; DB 7; Length 1898;
Best Local Similarity 100.0%; Pred. No. 4.7e-133;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 60
Db 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 74
QY 61 CATGTGAAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 75 CATGTGAAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 134
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 180
Db 135 AGCAGAAATCTGTATATATTTCCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 194
QY 181 CGGATACATACATATGATGATTTATATCTTGTTCGTGATTTGGCATCAGGACA 240
Db 195 CGGATACATACATATGATGATTTATATCTTGTTCGTGATTTGGCATCAGGACA 254
QY 241 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db 255 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 314

QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGAAATCAGTGTGG 360
Db |||||
QY 315 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGAAATCAGTGTGG 374
Db |||||
QY 361 CTTACACCAAGTTCTTACTGAGATGAATAAATAATTTGGATCCTAGTCTTTTATTGCTGAC 420
Db |||||
QY 375 CTTACACCAAGTTCTTACTGAGATGAATAAATAATTTGGATCCTAGTCTTTTATTGCTGAC 434
Db |||||
QY 421 TTTACAGCAACACGACGAGAGATTAGGATTATTA 453
Db |||||
QY 435 TTTACAGCAACACGACGAGAGATTAGGATTATTA 467
Db |||||

RESULT 8

US-10-472-533-175
; Sequence 175, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS906PCT
; CURRENT APPLICATION NUMBER: US/10/472,533
; CURRENT FILING DATE: 2003-09-20
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 650
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1398)..(1398)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1428)..(1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-472-533-175

Query Match 98.7%; Score 453; DB 9; Length 1898;
Best Local Similarity 100.0%; Pred. No. 4.7e-133;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCTTGATTGGACCGGTGCTGAGAACCTC 60
Db |||||
QY 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCTTGATTGGACCGGTGCTGAGAACCTC 74
Db |||||
QY 61 CATGTGAAATAAGTTGCTCTTGACTGTGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 120
Db |||||
QY 75 CATGTGAAATAAGTTGCTCTTGACTGTGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 134
Db |||||
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 180
Db |||||
QY 135 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 194
Db |||||
QY 181 CGGATACATACATATGTATATCTGATGTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 240
Db |||||
QY 195 CGGATACATACATATGTATATCTGATGTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 254
Db |||||
QY 241 AGGGTAGTTTCTGAGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db |||||
QY 255 AGGGTAGTTTCTGAGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 314
Db |||||
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db |||||
QY 315 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 374
Db |||||

QY 361 CTTACACCAAGTTCTTACTGAGATGAATAAATAATTTGGATCCTAGTCTTTTATTGCTGAC 420
Db |||||
QY 375 CTTACACCAAGTTCTTACTGAGATGAATAAATAATTTGGATCCTAGTCTTTTATTGCTGAC 434
Db |||||
QY 421 TTTACAGCAACACGACGAGAGATTAGGATTATTA 453
Db |||||
QY 435 TTTACAGCAACACGACGAGAGATTAGGATTATTA 467
Db |||||

RESULT 9

US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1704

Query Match 98.3%; Score 451.4; DB 6; Length 1492;
Best Local Similarity 99.8%; Pred. No. 1.3e-132;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCTTGATTGGACCGGTGCTGAGAACCTC 60
Db |||||
QY 52 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGTCTTGATTGGACCGGTGCTGAGAACCTC 111
Db |||||
QY 61 CATGTGAAATAAGTTGCTCTTGACTGTGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 120
Db |||||
QY 112 CATGTGAAATAAGTTGCTCTTGACTGTGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 171
Db |||||
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 180
Db |||||
QY 172 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAGAA 231
Db |||||
QY 181 CGGATACATACATATGTATATCTGATGTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 240
Db |||||
QY 232 CGGATACATACATATGTATATCTGATGTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 291
Db |||||
QY 241 AGGGTAGTTTCTGAGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db |||||
QY 292 AGGGTAGTTTCTGAGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 351
Db |||||
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db |||||
QY 352 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 411
Db |||||
QY 361 CTTACACCAAGTTTCTTACTGAGATGAATAAATAATTTGGATCCTAGTCTTTTATTGCTGAC 420
Db |||||
QY 412 CTTACACCAAGTTTCTTACTGAGATGAATAAATAATTTGGATCCTAGTCTTTTATTGCTGAC 471
Db |||||
QY 421 TTTACAGCAACACGACGAGAGATTAGGATTATTA 453
Db |||||
QY 472 TTTACAGCAACACGACGAGAGATTAGGATTATTA 504
Db |||||

RESULT 10

US-09-731-872-24
; Sequence 24, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste

APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78 US3 REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 24
LENGTH: 470
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 137..454
NAME/KEY: sig_peptide
LOCATION: 137..187
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 10.7019149919754
OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-09-731-872-24

Query Match 53.6%; Score 246; DB 3; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.1e-67;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 196
QY 61 CATGTGAAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 197 CATGTGAAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 256
QY 121 AGCAGAAATCTGTATATATTTGCGATGAATACATCTGGGAATGGGCTGCCCTGCAAA 180
Db 257 AGCAGAAATCTGTATATATTTGCGATGAATACATCTGGGAATGGGCTGCCCTGCAAA 316
QY 181 CGGATACATACATATGATATATGATTTATATATCTTGTTCGTGATTGGCATCAGGACA 240
Db 317 CGGATACATACATATGATATATGATTTATATCTTGTTCGTGATTGGCATCAGGACA 376
QY 241 AGGGTA 246
Db 377 AGGGTA 382

RESULT 11
US-09-876-997-24
Sequence 24, Application US/09876997
Publication No. US20030152921A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78 US4.CIP
CURRENT APPLICATION NUMBER: US/09/876,997
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 09/731,872
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 24
LENGTH: 470

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 137..454
NAME/KEY: sig_peptide
LOCATION: 137..187
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 10.7019149919754
OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-09-876-997-24
Query Match 53.6%; Score 246; DB 3; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.1e-67;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 60
Db 137 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTGCTGAGAACCTC 196
QY 61 CATGTGAAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 197 CATGTGAAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 256
QY 121 AGCAGAAATCTGTATATATTTGCGATGAATACATCTGGGAATGGGCTGCCCTGCAAA 180
Db 257 AGCAGAAATCTGTATATATTTGCGATGAATACATCTGGGAATGGGCTGCCCTGCAAA 316
QY 181 CGGATACATACATATGATATATGATTTATATATCTTGTTCGTGATTGGCATCAGGACA 240
Db 317 CGGATACATACATATGATATATGATTTATATCTTGTTCGTGATTGGCATCAGGACA 376
QY 241 AGGGTA 246
Db 377 AGGGTA 382

RESULT 12
US-10-643-836-24
Sequence 24, Application US/10643836
Publication No. US20050096458A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78 US3.REG
CURRENT APPLICATION NUMBER: US/10/643,836
CURRENT FILING DATE: 2003-08-19
PRIOR APPLICATION NUMBER: US/09/731,872
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 24
LENGTH: 470
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 137..454
NAME/KEY: sig_peptide
LOCATION: 137..187
OTHER INFORMATION: Von Heijne matrix
OTHER INFORMATION: score 10.7019149919754
OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-10-643-836-24
Query Match 53.6%; Score 246; DB 9; Length 470;

Best Local Similarity 100.0%; Pred. No. 2.1e-67;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGTTAGAAGTCTTGATGCTCCTCGCTGCTTGAATTTGGACCGGTGCTGAGAACCTC 60
Db |||||
Qy 137 ATGGCGTTAGAAGTCTTGATGCTCCTCGCTGCTTGAATTTGGACCGGTGCTGAGAACCTC 196
Db |||||

Qy 61 CATGTGAAAATAAGTCTCTCTGGACTGCTGATGCTCAGTTATCCAGTTGCAGAA 120
Db |||||

Qy 197 CATGTGAAAATAAGTCTCTCTGGACTGCTGATGCTCAGTTATCCAGTTGCAGAA 256
Db |||||

Qy 121 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 180
Db |||||

Qy 257 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 316
Db |||||

Qy 181 CGATACATACATATGTATATAGTTTATATATATCTTGGTATGTTGGCATCAGAGACA 240
Db |||||

Qy 317 CGATACATACATATGTATATAGTTTATATATATCTTGGTATGTTGGCATCAGAGACA 376
Db |||||

Qy 241 AGGGTA 246
Db |||||

Qy 377 AGGGTA 382
Db |||||

RESULT 13
US-09-984-429-532
; Sequence 532, Application US/09984429
; Publication No. US20040010132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P2
; CURRENT APPLICATION NUMBER: US/09/984,429
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,591
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 727
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 532
; LENGTH: 5013
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-429-532

Query Match 39.9%; Score 183; DB 3; Length 5013;
Best Local Similarity 100.0%; Pred. No. 8.6e-47;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 GTGAAAATAAGTTGCTCTCTGGACTGTTGATGCTCTCAGTTATCCAGTTGCAGAAAGC 123
Db |||||

Qy 124 AGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAATCGG 183
Db |||||

Qy 61 AGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAATCGG 120
Db |||||

Qy 184 ATACATACATATGTATATAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACAAGG 243
Db |||||

Db 121 ATACATACATATGTATATAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACAAGG 180
Qy 244 GTA 246
Db 181 GTA 183

RESULT 14
US-10-467-046-1
; Sequence 1, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 10115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2387...2501
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5443...5646
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6649...6747
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8907...9774
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2986)..(2986)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3847)..(3848)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3875)..(3879)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4105)..(4105)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4107)..(4109)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7347)..(7348)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8197)..(8197)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (8293)..(8293)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc.feature
LOCATION: (8334)..(8336)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc.feature
LOCATION: (8342)..(8349)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc.feature
LOCATION: (9942)..(9942)
OTHER INFORMATION: n is a, c, g, or t
US-10-467-046-1

Query Match 39.9%; Score 183; DB 8; Length 10115;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 64 GTGAAATAAGTTGCTCTCTGACTGGTTGATGGTCTCAGTTATCCAGTTTGCAGAAAGC 123
Db 5444 GTGAAATAAGTTGCTCTCTGACTGGTTGATGGTCTCAGTTATCCAGTTTGCAGAAAGC 5503
QY 124 AGAAATCTGTATATATTTGCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAAATCGG 183
Db 5504 AGAAATCTGTATATATTTGCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAAATCGG 5563
QY 184 ATACATACATATGTATATAGTTTATATATCTTGTTCGTGATGGCATCAGGACAAGG 243
Db 5564 ATACATACATATGTATATAGTTTATATATCTTGTTCGTGATGGCATCAGGACAAGG 5623
QY 244 GTA 246
Db 5624 GTA 5626

RESULT 15
US-10-287-971-39
Sequence 39, Application US/10287971
Publication No. US20040067882A1
GENERAL INFORMATION:
APPLICANT: Alsobrook, et al
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-480A
CURRENT APPLICATION NUMBER: US/10/287,971
PRIOR FILING DATE: 2002-11-05
PRIOR APPLICATION NUMBER: 09/997,425
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 10/035,568
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/338,626
PRIOR FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: 60/401,479
PRIOR FILING DATE: 2002-08-06
PRIOR APPLICATION NUMBER: 60/333,072
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: 60/348,283
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: 60/393,262
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 60/406,181
PRIOR FILING DATE: 2002-08-26
NUMBER OF SEQ ID NOS: 397
SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 39
LENGTH: 838
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (100)..(735)
US-10-287-971-39

Query Match 19.7%; Score 90.2; DB 7; Length 838;
Best Local Similarity 57.7%; Pred. No. 1.3e-17;
Matches 161; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 8 TAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTGGACCGGTGCTCAGAACCTCCATGTGA 67
Db 119 TAGGACTGATGATCCTCCTCACCTCTGCGCTTTTCAGCCGGTTTCAGGACAAAGTCCAATGA 178
QY 68 AAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTTGCAGAAAGCAGAA 127
Db 179 CTGTGCTGTCTCCATAGACTGGTTTCATGGTCAAGTGCACCCCTTCATGCTTAACAACG 238
QY 128 ATCTGTATATATTTGCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATAC 187
Db 239 ATGTGTGTGTACACTTTTCATGAACCTACACTTGGGCTGGGTTGCCCCCAAAACCATGTTTC 298
QY 248 TTTCTGAGGAACCTCCTCTTTTCAAACCGAGCTGTACT 286
Db 359 TCTCTCAGGACATGGTTATCTACAGCACTGAGATACACT 397

Search completed: May 9, 2006, 15:02:51
Job time : 1222.4 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: May 9, 2006, 13:13:06 ; Search time 945 Seconds
(without alignments)
1978.089 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511
Perfect score: 459
Sequence: 1 atggcgtagaagtctgat.....agttagattattaatttc 459

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Databases : Published Applications NA New:*
1: /SIDSS/ptodata/1/pubpna/US08 NEW PUB.seq1:*
2: /SIDSS/ptodata/1/pubpna/US06 NEW PUB.seq2:*
3: /SIDSS/ptodata/1/pubpna/US07 NEW PUB.seq3:*
4: /SIDSS/ptodata/1/pubpna/US08 NEW PUB.seq4:*
5: /SIDSS/ptodata/1/pubpna/PCT NEW PUB.seq5:*
6: /SIDSS/ptodata/1/pubpna/US03 NEW PUB.seq6:*
7: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq7:*
8: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq8:*
9: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq9:*
10: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq10:*
11: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq11:*
12: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq12:*
13: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq13:*
14: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq14:*
15: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq15:*
16: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq16:*
17: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq17:*
18: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq18:*
19: /SIDSS/ptodata/1/pubpna/US60_NEW_PUB.seq19:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	451.4	98.3	1492	18	US-11-072-512-1704
2	446	97.2	497	11	US-10-475-075-564
3	429	93.5	500	11	US-10-475-075-563
4	90.2	19.7	1141	10	US-10-821-234-181
C 5	39.4	8.6	564	7	US-09-925-065A-768295
C 6	39.4	8.6	580	7	US-09-925-065A-46080
C 7	39.4	8.6	580	7	US-09-925-065A-46081
C 8	39.4	8.6	580	11	US-10-301-480-147318
C 9	39.4	8.6	580	11	US-10-301-480-147319
C 10	39.4	8.6	580	12	US-10-301-480-760727
C 11	39.4	8.6	580	12	US-10-301-480-760728
C 12	39.4	8.6	1524	7	US-09-925-065A-29822
C 13	39.4	8.6	1524	7	US-09-925-065A-29823
C 14	39.4	8.6	1524	7	US-09-925-065A-47775

C 15	39.4	8.6	1524	7	US-09-925-065A-47776
C 16	39.4	8.6	1524	11	US-10-301-480-111060
C 17	39.4	8.6	1524	11	US-10-301-480-131061
C 18	39.4	8.6	1524	11	US-10-301-480-149013
C 19	39.4	8.6	1524	11	US-10-301-480-149014
C 20	39.4	8.6	1524	12	US-10-301-480-744469
C 21	39.4	8.6	1524	12	US-10-301-480-744470
C 22	39.4	8.6	1524	12	US-10-301-480-762422
C 23	39.4	8.6	1524	12	US-10-301-480-762423
C 24	37.4	8.1	199130	10	US-10-995-561-13233
C 25	35.4	7.7	579	7	US-09-925-065A-276453
C 26	35.2	7.7	387780	10	US-10-995-561-13259
C 27	34.6	7.5	835	12	US-10-301-480-548415
C 28	34.6	7.5	835	12	US-10-301-480-1161824
C 29	34.4	7.5	1111	10	US-10-750-185-54368
C 30	34.4	7.5	1111	10	US-10-750-623-54368
C 31	34.2	7.5	1221	7	US-09-925-065A-546990
C 32	34.2	7.5	1221	12	US-10-301-480-524940
C 33	34.2	7.5	1221	12	US-10-301-480-1138349
C 34	34	7.4	726	7	US-09-925-065A-940832
C 35	34	7.4	1088	7	US-09-925-065A-552224
C 36	34	7.4	1088	12	US-10-301-480-530174
C 37	34	7.4	1088	12	US-10-301-480-1143583
C 38	34	7.4	159497	17	US-11-112-908-61
C 39	34	7.4	171427	17	US-11-112-908-60
C 40	33.8	7.4	1241	11	US-10-301-480-37618
C 41	33.8	7.4	1241	11	US-10-301-480-37619
C 42	33.8	7.4	1241	12	US-10-301-480-651027
C 43	33.8	7.4	1241	12	US-10-301-480-651028
C 44	33.6	7.3	608	7	US-09-925-065A-506156
C 45	33.6	7.3	717	12	US-10-301-480-612896

ALIGNMENTS

RESULT 1
US-11-072-512-1704
; Sequence 1704, Application US/11072512
; Publication No. US2006029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens

US-11-072-512-1704

Query Match	98.3%	Score 451.4	DB 18	Length 1492
Best Local Similarity	99.8%	Pred. No. 1.3e-110		
Matches 452	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	ATGCGTTAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTTGGACCGGTGCTGAGAACCTC	60	
Db	52	ATGCGTTAGAAGTCTTGATGCTCCTCGCTGCTCTTGATTTGGACCGGTGCTGAGAACCTC	111	
Qy	61	CATGTGAAATAAGTTGCTCTCTGGACTGGTGTGATGCTCAGTTATCCGAGTTGCAGAA	120	
Db	112	CATGTGAAATAAGTTGCTCTCTGGACTGGTGTGATGCTCAGTTATCCGAGTTGCAGAA	171	
Qy	121	AGCAGAAATCTGTATATATTTGCGGATGCAATTACATCTGGGAATGGGTGCCTGCAAA	180	
Db	172	AGCAGAAATCTGTATATATTTGCGGATGCAATTACATCTGGGAATGGGTGCCTGCAAA	231	
Qy	181	CGGATACATACATATGATATAGTATTATATATCTTGTTCTGTATGTGGCATCAGGACA	240	
Db	232	CGGATACATACATATGATATAGTATTATATATCTTGTTCTGTATGTGGCATCAGGACA	291	
Qy	241	AGGTTAGTTTCTGAGGAAACTCTCCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT	300	
Db	292	AGGTTAGTTTCTGAGGAAACTCTCCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT	351	
Qy	301	ATAGATCATGACCTTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCATGTGFG	360	
Db	352	ATAGATCATGACCTTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCATGTGFG	411	
Qy	361	CTTACACCGTTTCTACTGAGAAATGAATAAAATTTGGATCCTAGTCTTTATTGCTGCAC	420	
Db	412	CTTACACCGTTTCTACTGAGAAATGAATAAAATTTGGATCCTAGTCTTTATTGCTGCAC	471	
Qy	421	TTTTCAGACACAGCAGAGAGTATTAGGATTATTA	453	
Db	472	TTTTCAGACACAGCAGAGAGTATTAGGATTATTA	504	

```

RESULT 2
US-10-475-075-564
; Sequence 564, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 564
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 52...495
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 52..102
; OTHER INFORMATION: Von Heinje matrix
; OTHER INFORMATION: score 10.6999998092651
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-10-475-075-564

```

Query Match 97.2%; Score 446; DB 11; Length 497;

Best Local Similarity 100.0%; Pred. No. 2.7e-109;		
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGCGGTTAGAAGCTTTGATGCTCTCGCTGTCCTTGATTTGGACCGGTGCTGAGAACCTC 60
Db	52	ATGCGGTTAGAAGCTTTGATGCTCTCGCTGTCCTTGATTTGGACCGGTGCTGAGAACCTC 111
Qy	61	CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTATTCCAGTTGCGAGAA 120
Db	112	CATGTGAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTATTCCAGTTGCGAGAA 171
Qy	121	AGCAGAAATCTGTATATATATTTCGGATGAATTACATCTGGGAATGGCTGCCCTGCCAAAT 180
Db	172	AGCAGAAATCTGTATATATTTCGGATGAATTACATCTGGGAATGGCTGCCCTGCCAAAT 231
Qy	181	CGGATACATACATATGTATATAGTTTATATATATCTGTTCTGTGATTTGGCATCAGGACA 240
Db	232	CGGATACATACATATGTATATAGTTTATATATATCTGTTCTGTGATTTGGCATCAGGACA 291
Qy	241	AGGGTAGTTTCTGAGGAAACTCTCTCTTTTTCAAAACCGAGCTGTACTTTTACCCCAAGGAAT 300
Db	292	AGGGTAGTTTCTGAGGAAACTCTCTCTTTTTCAAAACCGAGCTGTACTTTTACCCCAAGGAAT 351
Qy	301	ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCCACTCTAGGAAATCAGTGTGG 360
Db	352	ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCCACTCTAGGAAATCAGTGTGG 411
Qy	361	CTTTACACCAAGTTTCTACTGAGAATGAAATAAAATTTGGATCCTAGTCTCTTTTATTGCTGAC 420
Db	412	CTTTACACCAAGTTTCTACTGAGAATGAAATAAAATTTGGATCCTAGTCTCTTTTATTGCTGAC 471
Qy	421	TTTTCAGACAACAGCAGAGAAGATTAGG 446
Db	472	TTTTCAGACAACAGCAGAGAAGATTAGG 497

```

RESULT 3
US-10-475-075-563
; Sequence 563, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 563
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 72..500
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 72..122
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLAVLIWTGA/EN
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1..3
; OTHER INFORMATION: n = a, g, c or t
US-10-475-075-563

```

Query Match	93.5%;	Score 429;	DB 11;	Length 500;
Best Local Similarity	100.0%;	Pred. No. 9.7e-105;		
Matches 429;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCTGAGAACCTC	60		
Db				
Qy 72	ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCTGAGAACCTC	131		
Db				
Qy 61	CATGTGAAAAAAGTTGCTCTCTGACCTGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA	120		
Db				
Qy 132	CATGTGAAAAAAGTTGCTCTCTGACCTGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA	191		
Db				
Qy 121	AGCAGAAATCTGTATATATTTGCGGATGAAATPACATCTGGGAATGGCTGCCCTGCAAA	180		
Db				
Qy 192	AGCAGAAATCTGTATATATTTGCGGATGAAATPACATCTGGGAATGGCTGCCCTGCAAA	251		
Db				
Qy 181	CGGATACATACATATGATATAGAGTTTATATATCTTGTTCTGTGATTGGCATCAGGACA	240		
Db				
Qy 252	CGGATACATACATATGATATAGAGTTTATATATCTTGTTCTGTGATTGGCATCAGGACA	311		
Db				
Qy 241	AGGGTAGTTTCTGAGGAAATCTCCTTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT	300		
Db				
Qy 312	AGGGTAGTTTCTGAGGAAATCTCCTTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT	371		
Db				
Qy 301	ATAGATCATGACCCCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAAATCAGTGTGG	360		
Db				
Qy 372	ATAGATCATGACCCCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAAATCAGTGTGG	431		
Db				
Qy 361	CTTACACGAGTTTCTACTGAGAAATGAAATAAAAATGGATCCTAGTCTCTTTTATTGCTGAC	420		
Db				
Qy 432	CTTACACGAGTTTCTACTGAGAAATGAAATAAAAATGGATCCTAGTCTCTTTTATTGCTGAC	491		
Db				
Qy 421	TTTTCAGACA	429		
Db				
Qy 492	TTTTCAGACA	500		
Db				

```

RESULT 4
US-10-821-234-181
; Sequence 181, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Suesan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 181
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-181

Query Match          19.7%;   Score 90.2;   DB 10;   Length 1141;
Best Local Similarity 57.7%;   Pred No. 4, 7e-14;
Matches 161;   Conservative 0;   Mismatches 118;   Indels 0;   Gaps 0

QY      8   TAGAGCTCTTGANGCTCTCGCTGCTCTTCGATTGATTGGACCGGTGCTGAGAACCTCCATGTGA 67
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      311 TAGGACTGATGATCTCTCTCACTCTCGGTTTTTCAGCCGGTTTCAGGACAAAGTCCAATGA 370

QY      68   AATAAGTGTGCTCTCGGACTGGTTGATGCTCAGTTATCCCAAGTTGCAGAAAGCAGAA 127
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      371   CTGTGCTGTGCTCCATAGACGTGGTTTCATGTGTCACAGTCGACCCCTTCATGCTAAACAACG 430

QY      128  ATCTGTATATATTTTCGGGATGAATTCATCATCTCGGAAATGGCTGCCTGCAAAATCGGATAC 187

```

Db	431	ATGTGTGTACACTTTTCATGAACCTACACTTGGGCTGGTTGCCCCCAACCATGTC	490
Qy	198	ATACATATGTATAGAGTTTATATATCTGTTCGTGATTTGGCATCAGGCAAGGGTAG	247
Db	491	AGCCACAGCCCTACCAGTTCCACTACCGTGTCTACTGAATGTGGCATCAGGGCCAAAGCTG	550
Qy	248	TTTCTGAGAACTCTCCCTTTTCAAACCGAGCTGTACT	286
Db	551	TCTCTCAGGACATGGTTATCTACAGCACTGAGATACACT	589

RESULT 5

US-09-925-065A-768295/c

; Sequence 768295, Application US/09925065A

; Publication No. US20040181048A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single

; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.135

; CURRENT APPLICATION NUMBER: US/09/925.065A

; CURRENT FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147

; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/289,846

; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 768295

; LENGTH: 564

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-925-065A-768295

Query Match 8.6%; Score 39.4; DB 7; Length 564;

Best Local Similarity 53.6%; Pred. No. 1.5;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0

Qy	91	TTGATGCTCTCAGTTATCCAGTTGCGAGAAGCAGAAATCTGTATATATTTTGGCGATGAA	150
Db	307	TTGATGAATGAGCATGCAGATATGTATATAGCATATATCTACATTTTATTTAAAAATACA	248
Qy	151	TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATACATATGTATAGCTTTTATA	210
Db	247	TGCCAAATATGCATAATATATATATATCCATGACATAAATGCATATATTTATATATAAA	188
Qy	211	TATCTTCTTCGTGATTTGGCATCAGGACAAGG	243
Db	187	TCTATTGGAAGAAATTTGTATACATAAGGATGTGG	155

RESULT 6

US-09-925-065A-46080/c

; Sequence 46080, Application US/09925065A

; Publication No. US20040181048A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single

; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.135

; CURRENT APPLICATION NUMBER: US/09/925.065A

; CURRENT FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147

; PRIOR FILING DATE: 2000-11-20

```
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46080
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-46080

Query Match      8.6%; Score 39.4; DB 7; Length 580;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAAATGAGCATGCAGATATGTATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGCTGCCCTGCAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAATATATATATATCCAAATGCATAATATTTATATATATAAA 144

QY 211 TATCTTGTTCGATTGTGGCATCAGGCAAGG 243
Db 143 TCTATTGGAAGAAATTTTACATAAGGATGG 111

RESULT 7
US-09-925-065A-46081/c
; Sequence 46081, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46081
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-46081

Query Match      8.6%; Score 39.4; DB 7; Length 580;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAAATGAGCATGCAGATATGTATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGCTGCCCTGCAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAATATATATATATCCAAATGCATAATATTTATATATATAAA 144

QY 211 TATCTTGTTCGATTGTGGCATCAGGCAAGG 243

US-10-301-480-147318/c
; Sequence 147318, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147318
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-147318

Query Match      8.6%; Score 39.4; DB 11; Length 580;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAAATGAGCATGCAGATATGTATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGCTGCCCTGCAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAATATATATATATCCAAATGCATAATATTTATATATATAAA 144

QY 211 TATCTTGTTCGATTGTGGCATCAGGCAAGG 243
Db 143 TCTATTGGAAGAAATTTTACATAAGGATGG 111

RESULT 9
US-10-301-480-147319/c
; Sequence 147319, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147319
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-147319

Query Match      8.6%; Score 39.4; DB 11; Length 580;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAAATGAGCATGCAGATATGTATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGCTGCCCTGCAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAATATATATATATCCAAATGCATAATATTTATATATATAAA 144

QY 211 TATCTTGTTCGATTGTGGCATCAGGCAAGG 243
```



```
Db 263 TTGATGAATGAGCATGCAGATATGTATATAGCATATATCTACATTTATTTATAAATAACA 204
Qy 151 TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATACATATGATGATTTATA 210
Db 203 TGCCAAATATGCATAAATATATATATCCATGACATAAATGCATATATTTATATATATAA 144
Qy 211 TATCTTGTCTGATTTGGCATCAGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111

RESULT 10
US-10-301-480-760727/c
; Sequence 760727, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 760727
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-760727

Query Match 8.6%; Score 39.4; DB 12; Length 580;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAATGAGCATGCAGATATGTATATAGCATATATCTACATTTATTTATAAATAACA 204
Qy 151 TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATACATATGATGATTTATA 210
Db 203 TGCCAAATATGCATAAATATATATATCCATGACATAAATGCATATATTTATATATAA 144
Qy 211 TATCTTGTCTGATTTGGCATCAGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111

RESULT 11
US-10-301-480-760728/c
; Sequence 760728, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 760728
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-760728

Query Match 8.6%; Score 39.4; DB 12; Length 580;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAATGAGCATGCAGATATGTATATAGCATATATCTACATTTATTTATAAATAACA 204
Qy 151 TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATACATATGATGATTTATA 210
Db 203 TGCCAAATATGCATAAATATATATATCCATGACATAAATGCATATATTTATATATAA 144
Qy 211 TATCTTGTCTGATTTGGCATCAGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111

RESULT 12
US-09-925-065A-29822/c
; Sequence 29822, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29822
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-29822

Query Match 8.6%; Score 39.4; DB 7; Length 1524;
Best Local Similarity 53.6%; Pred. No. 2;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 91 TTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATGAA 150
Db 263 TTGATGAATGAGCATGCAGATATGTATATAGCATATATCTACATTTATTTATAAATAACA 204
Qy 151 TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATACATATGATGATTTATA 210
Db 203 TGCCAAATATGCATAAATATATATATCCATGACATAAATGCATATATTTATATATAA 144
Qy 211 TATCTTGTCTGATTTGGCATCAGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111

RESULT 13
US-09-925-065A-29823/c
; Sequence 29823, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
```

```
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29823
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-29823

Query Match      8.6%; Score 39.4; DB 7; Length 1524;
Best Local Similarity 53.6%; Pred. No. 2;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTCGAGAAAGCAGAAATCTGTATATATTTTCGGATGAA 150
Db 263 TTGATGAATGAGCAGATGATATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGGCTGCCGAAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAATATATATATCCAAATGCATAAATGCATATATTTTATATATAAA 144

QY 211 TATCTTGTTCGATGTGGCATCAGGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111
```

```
RESULT 14
US-09-925-065A-47775/c
; Sequence 47775, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47775
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-47775

Query Match      8.6%; Score 39.4; DB 7; Length 1524;
Best Local Similarity 53.6%; Pred. No. 2;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTCGAGAAAGCAGAAATCTGTATATATTTTCGGATGAA 150
Db 263 TTGATGAATGAGCAGATGATATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGGCTGCCGAAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAATATATATATCCAAATGCATAAATGCATATATTTTATATATAAA 144

QY 211 TATCTTGTTCGATGTGGCATCAGGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111
```

```
QY 151 TTACATCTGGGAATGGGCTGCCGAAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAAATATATATATCCAAATGCATAAATGCATATATTTTATATATAAA 144

QY 211 TATCTTGTTCGATGTGGCATCAGGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111

RESULT 15
US-09-925-065A-47776/c
; Sequence 47776, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47776
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-47776

Query Match      8.6%; Score 39.4; DB 7; Length 1524;
Best Local Similarity 53.6%; Pred. No. 2;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 91 TTGATGGTCTCAGTTATCCAGTTCGAGAAAGCAGAAATCTGTATATATTTTCGGATGAA 150
Db 263 TTGATGAATGAGCAGATGATATATAGCATATATCTACATTTTATTAAAAATACA 204

QY 151 TTACATCTGGGAATGGGCTGCCGAAATCGGATACATACATATGTATATGAGTTTATA 210
Db 203 TGCCAAATATGCATAAATATATATATCCAAATGCATAAATGCATATATTTTATATATAAA 144

QY 211 TATCTTGTTCGATGTGGCATCAGGACAAGG 243
Db 143 TCTATTGGAAGAAATTTGTACATAAGGATGTGG 111

Search completed: May 9, 2006, 13:33:06
Job time : 946 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2006, 12:54:43 ; Search time 5670 Seconds
(without alignments)
3787.526 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511
Perfect score: 459
Sequence: 1 atggcgttagaagctctgat.....agttagattattataatttc 459

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_hic.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_est7.*
- 9: gb_gss1.*
- 10: gb_gss2.*
- 11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453	98.7	626	BI463012	603204535
2	453	98.7	776	BI463012	603204535
3	453	98.7	878	BI520154	603071295
4	440	95.9	577	BI118111	BI118111
5	418.8	91.2	778	BI459538	603200545
6	410.4	89.4	619	AW664990	h198h06.x
7	378.6	82.5	516	BE972717	601652136
8	318	69.3	508	AW572898	h17c01.x
9	303	66.0	430	AA608835	af03h06.s
10	280	63.2	480	AI554902	tes3903.x
11	246	53.6	331	AI138881	qd99c04.x
12	236.2	51.5	598	BO801002	H3059H12-
13	236.2	51.5	629	BY731676	BY731676
14	234.6	51.1	862	CO810251	AGENCOURT
15	219	47.7	642	BB557840	BB557840
16	187	40.7	321	AI028204	ov96b07.x
17	171	37.3	496	BO67944	H3059H12-
18	161.8	35.3	587	C88054	C88054
19	161	35.1	504	CO798599	AGENCOURT
20	151.8	33.1	576	AU045772	AU045772
21	127.4	27.8	168	CB298889	220023.re
22	113.2	24.7	688	CO799574	AGENCOURT

C 23	108.2	23.6	389	1	AU044924	AU044924
C 24	101	22.0	559	9	AQ723272	HS_2085_B
C 25	90.2	19.7	366	2	BG184286	RST3209_A
26	90.2	19.7	583	3	BP321198	BP321198
27	90.2	19.7	627	2	BG620477	BG620477
28	90.2	19.7	666	2	BG620807	BG620807
29	90.2	19.7	696	6	CB996101	AGENCOURT
30	90.2	19.7	820	6	CF593309	AGENCOURT
31	90.2	19.7	857	2	BG623009	602647871
32	90.2	19.7	882	2	BG199300	RST18581
33	90.2	19.7	883	5	BX386304	BX386304
34	90.2	19.7	895	6	CA489318	AGENCOURT
35	90.2	19.7	943	4	CR604394	full-leng
36	90.2	19.7	965	5	BX359837	BX359837
37	90.2	19.7	1056	4	CR618096	full-leng
38	90.2	19.7	1075	4	CR603619	full-leng
C 39	89.8	19.6	942	5	BX356611	BX356611
40	89.2	19.4	1044	5	BX361440	BX361440
41	88.6	19.3	752	5	EX110037	EX110037
42	88.6	19.3	812	2	BG202365	RST21722
43	88.6	19.3	860	2	BG203907	RST23299
44	88.2	19.2	624	8	DR004851	TC112894
C 45	87.8	19.1	949	5	BX386303	BX386303

ALIGNMENTS

BI463012 626 bp mRNA linear EST 21-AUG-2001
603204535F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5270384 5',
mRNA sequence.
BI463012
BI463012.1 GI:15253668
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 626)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1682 row: e column: 09
High quality sequence stop: 624.
Location/Qualifiers
1. 626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5270384"
/lab_host="DH10B"
/clone_lib="NIH MGC 97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI
(gtcag); Oligo-dt primer using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein

(NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

```
ORIGIN
Query Match      98.7%; Score 453; DB 3; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e-119;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTGTTGATTTGGACCGGTGCTGAGAACCTC 60
DB 57 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTGTTGATTTGGACCGGTGCTGAGAACCTC 116
QY 61 CATGTGAAATAAAGTTGCTCTCTGACCTGGTTGATGCTCTCAGTTATCCAGTTGCAGAA 120
DB 117 CATGTGAAATAAAGTTGCTCTCTGACCTGGTTGATGCTCTCAGTTATCCAGTTGCAGAA 176
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 177 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 236
QY 181 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 237 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296
QY 241 AGGGTAGTTTCTGAGGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
DB 297 AGGGTAGTTTCTGAGGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 356
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCAGGAAATCAGTGTGG 360
DB 357 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCAGGAAATCAGTGTGG 416
QY 361 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCTTCTGATGCTGCTTTTATTCCTGAC 420
DB 417 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCTTCTGATGCTGCTTTTATTCCTGAC 476
QY 421 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 453
DB 477 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 509

RESULT 2
BG772527      776 bp      mRNA      linear      EST 15-MAY-2001
LOCUS      602720619F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837612 5',
DEFINITION      mRNA sequence.
ACCESSION      BG772527
VERSION      BG772527.1 GI:14083180
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1 (bases 1 to 776)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNALS      Unpublished (1999)
COMMENTS      Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLNL0770 row: e column: 05
High quality sequence stop: 773.
Location/Qualifiers
1..776
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4837612"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."
```

ORIGIN

```
Query Match      98.7%; Score 453; DB 2; Length 776;
Best Local Similarity 100.0%; Pred. No. 1.4e-119;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTGTTGATTTGGACCGGTGCTGAGAACCTC 60
DB 57 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTGTTGATTTGGACCGGTGCTGAGAACCTC 116
QY 61 CATGTGAAATAAAGTTGCTCTCTGACCTGGTTGATGCTCTCAGTTATCCAGTTGCAGAA 120
DB 117 CATGTGAAATAAAGTTGCTCTCTGACCTGGTTGATGCTCTCAGTTATCCAGTTGCAGAA 176
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 177 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAAAT 236
QY 181 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 237 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296
QY 241 AGGGTAGTTTCTGAGGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
DB 297 AGGGTAGTTTCTGAGGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 356
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCAGGAAATCAGTGTGG 360
DB 357 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCAGGAAATCAGTGTGG 416
QY 361 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCTTCTGATGCTGCTTTTATTCCTGAC 420
DB 417 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCTTCTGATGCTGCTTTTATTCCTGAC 476
QY 421 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 453
DB 477 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 509
```

RESULT 3

```
BG772527      878 bp      mRNA      linear      EST 29-AUG-2001
LOCUS      603071295F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163455 5',
DEFINITION      mRNA sequence.
ACCESSION      BG772527
VERSION      BG772527.1 GI:15344946
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1 (bases 1 to 878)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNALS      Unpublished (1999)
COMMENTS      Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
```

FEATURES
source


```
Best Local Similarity 98.2%; Pred. No. 2.7e-107;
Matches 425; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 21 GCTCCTCGCTGTTGATTTGGACCGGTGCTGAGAACTCTCATGTGAAAATAAGTTGCTC 80
Db 614 GGTCTCGCTGCTGTGATTT-GACCGGTGCTGAGACCTCCATGTGAAAATAAGTTGCTC 556
QY 81 TCTGGACTGCTGTGATGCTCTCAGTTATCCAGTTTGCAAGAACGAGAAATCTGTATATTT 140
Db 555 TCTGACTTGGTGTGATGCTCTCAGTTATCCAGTTTGCAAGAACGAGAAATCTGTATATTT 496
QY 141 TGGGATGAATACATCTGGAATGGCTGCCCTGCAAAATCGGATACATACATATGTATA 200
Db 495 TGGGATGAATACATCTGGAATGGCTGCCCTGCAAAATCGGATACATACATATGTATA 436
QY 201 TGAGTTTATATATCTTGTTCGTGATGTGTCATCAGCAAGGGTAGTTTCTGAGGAAC 260
Db 435 TGAGTTTATATATCTTGTTCGTGATGTGTCATCAGCAAGGGTAGTTTCTGAGGAAC 376
QY 261 TCTCCTTTTCAAAACCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCTCAGGA 320
Db 375 TCTCCTTTTCAAAACCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCTCAGGA 316
QY 321 AATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGGCTTACACCACTTCTACTGA 380
Db 315 AATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGGCTTACACCACTTCTACTGA 256
QY 381 GAATGAATAAATTCGATCCTAGTCTTATTTATTTGCTGACTTTCAGACACAGCAAGA 440
Db 255 GAATGAATAAATTCGATCCTAGTCTTATTTATTTGCTGACTTTCAGACACAGCAAGA 196
QY 441 GTTAGGATTATTA 453
Db 195 GTTAGGATTATTA 183

RESULT 7
BE972717 516 bp mRNA linear EST 04-OCT-2000
LOCUS 601652136F1 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:3935288 5',
DEFINITION mRNA sequence.
ACCESSION BE972717
VERSION BE972717.1 GI:10586053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW777 row: h column: 09
High quality sequence stop: 505.
Location/Qualifiers
1. .516
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3935288"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC 82"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:

sfII (ggccgctcgcc); Site 2: sfII (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGGCGCCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

ORIGIN

Query Match 82.5%; Score 378.6; DB 2; Length 516;
Best Local Similarity 98.7%; Pred. No. 4e-98; Indels 1; Gaps 1;
Matches 392; Conservative 0; Mismatches 4;

QY 1 ATGGCGTTAGAACTTTGATGCTCTCGCTGCTCTTGAATTTGGACCGGTGCTGAGAACTC 60
Db 120 ATGGCGTTAGAACTTTGATGCTCTCGCTGCTCTTGAATTTGGACCGGTGCTGAGAACTC 179
QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAAG 120
Db 180 CATGTGAAAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAAG 239
QY 121 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
Db 240 AGCAGAAATCTGTATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 299
QY 181 CGGATACATACATATATATATATATATATATATCTTGTGTTGTCATGATGTCATCAGGACA 240
Db 300 CGGATACATACATATATATATATATATATATATCTTGTGTTGTCATGATGTCATCAGGACA 359
QY 241 AGGTTAGTTTCTGAGAAACTCTCTTTTCAAACCGAGCTGCTTATCCCAAGGAA- 299
Db 360 AGGTTAGTTTCTGAGAAACTCTCTTTTCAAACCGAGCTGCTTATCCCAAGGAA- 419
QY 300 TATAGATCATGACCTCAGGAATCCATTTGGAGTGTCCACCTTAGGAAATCAGTGTG 359
Db 420 TATAGATCATGACCTCAGGAATCCATTTGGAGTGTCCACCTTAGGAAATCAGTGTG 479
QY 360 GCTTACACCAAGTTTCTACTGAGAAATGAAAATAAATG 396
Db 480 GCTTACACCAAGTTTCTACTGAGAAATGAAAATAAATG 516

RESULT 8
AW572898/c 508 bp mRNA linear EST 13-MAR-2000
LOCUS hf17c01.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:2932128 3', mRNA sequence.
ACCESSION AW572898
VERSION AW572898.1 GI:7237631
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 508)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 451.
Location/Qualifiers
1. .508
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:2932128"
 /lab_host="DH10B"
 /clone_lib="Soares NFL T GBC S1"
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI CGAP GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 69.3%; Score 318; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 1.3e-80;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 TGCTTTGATTTCGACGGTCTGAGAACCTCCATGTGAAATAAGTTGCTCTCGACTG 89
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 508 TGCTTTGATTTCGACGGTCTGAGAACCTCCATGTGAAATAAGTTGCTCTCGACTG 449

QY 90 GTTGATGCTCTCAGTTATCCCATGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATCA 149
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 448 GTTGATGCTCTCAGTTATCCCATGTTGCAGAAAGCAGAAATCTGTATATATTTGCGGATCA 389

QY 150 ATTACATCTGGGAATGGGCTGCCCTGCGAAATCGGATACATCATATGATATGAGTTTAT 209
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 388 ATTACATCTGGGAATGGGCTGCCCTGCGAAATCGGATACATCATATGATATGAGTTTAT 329

QY 210 ATATCTTGTCGTGATTGGCATCAGGACAGGGTAGTTTCTGAGGAACTCTCCTTTT 269
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 328 ATATCTTGTCGTGATTGGCATCAGGACAGGGTAGTTTCTGAGGAACTCTCCTTTT 269

QY 270 TCAACCCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCTCAGGAAATCCATTT 329
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 268 TCAACCCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCTCAGGAAATCCATTT 209

QY 330 GGAGTGTTCACCTCTAG 347
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 208 GGAGTGTTCACCTCTAG 191

RESULT 9

AA608835/c
 LOCUS af03h06.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1030619
 DEFINITION 3', mRNA sequence.
 ACCESSION AA608835
 VERSION AA608835.1 GI:2457263
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 430)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 629 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 424.

FEATURES

source
 Location/Qualifiers
 1..430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1030619"
 /sex="male"
 /lab_host="DH10B"
 /clone_lib="Soares testis_NHT"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGGAGCGGCCCAATTTTTTTTTTTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 66.0%; Score 303; DB 1; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2.7e-76;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATATGATATGAGTTTATA 210
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 430 TTACATCTGGGAATGGGCTGCCCTGCAAAATCGGATACATATGATATGAGTTTATA 371

QY 211 TATCTTGTTCGTGATTGGCATCAGGACAGGGTAGTTTCTGAGGAACTCTCCTTTT 270
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 370 TATCTTGTTCGTGATTGGCATCAGGACAGGGTAGTTTCTGAGGAACTCTCCTTTT 311

QY 271 CAACCCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCTCAGGAAATCCATTTG 330
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 310 CAACCCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCTCAGGAAATCCATTTG 251

QY 331 GAGTGTTCACCTCTAGGAAATCAGTGTGGCTTACCCAGTTTCTACTGAGAATAAATA 390
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 250 GAGTGTTCACCTCTAGGAAATCAGTGTGGCTTACCCAGTTTCTACTGAGAATAAATA 191

QY 391 AAATGGATCTAGTCTCTTTTATGCTGACTTTCAGACAAACAGCAGAGCTTAGGATTA 450
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 190 AAATGGATCTAGTCTCTTTTATGCTGACTTTCAGACAAACAGCAGAGCTTAGGATTA 131

QY 451 TTA 453
 Db |||||
 130 TTA 128

RESULT 10

AI554902/c
 LOCUS tes3g03.x1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:2090452 3', mRNA sequence.
 ACCESSION AI554902
 VERSION AI554902.1 GI:4487265
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 480)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 656 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 458.
Location/Qualifiers
1. .480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2090452"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: pooled; Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES

source

ORIGIN

Query Match 63.2%; Score 290; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.6e-72;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 CTCATGTGAAATAAGTTGCTCTCTGGACTGGTGTGATGCTTCAGTTATCCAGTTGCA 117
Db 480 CTCATGTGAAATAAGTTGCTCTCTGGACTGGTGTGATGCTTCAGTTATCCAGTTGCA 421
QY 118 GAAAGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGCTGCCCTGCA 177
Db 420 GAAAGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGCTGCCCTGCA 361
QY 178 AATCGGATACATACATATGTATAGTTTATATATCTTGTTCGTGATTGTGCATCAGG 237
Db 360 AATCGGATACATACATATGTATAGTTTATATATCTTGTTCGTGATTGTGCATCAGG 301
QY 238 ACAAGGGTAGTTTCTCAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTACCCCAAGG 297
Db 300 ACAAGGGTAGTTTCTCAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTACCCCAAGG 241
QY 298 AATATAGATCATGACCTCAGGAATCCATTGGAGTGTCCACCTCTAG 347
Db 240 AATATAGATCATGACCTCAGGAATCCATTGGAGTGTCCACCTCTAG 191

RESULT 11

Al138881/c 331 bp mRNA linear EST 28-OCT-1998
LOCUS qd99c04.x1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1737606
DEFINITION 3', mRNA sequence.
ACCESSION Al138881
VERSION Al138881.1 GI:3644853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 331)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 395 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES

source

1. .331
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1737606"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/notes="Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', TGTACCATCTGAAGTGGAGCGCGGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 53.6%; Score 246; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 7.4e-60;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGTGTGAGAACCTC 60
Db 331 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGTGTGAGAACCTC 272
QY 61 CATGTGAAATTAAGTTGCTCTCTGGACTGGTGTGATGCTTCAGTTATCCAGTTGAGAA 120
Db 271 CATGTGAAATTAAGTTGCTCTCTGGACTGGTGTGATGCTTCAGTTATCCAGTTGAGAA 212
QY 121 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTCGGAATGGCTGCCCTGCAAA 180
Db 211 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTCGGAATGGCTGCCCTGCAAA 152
QY 181 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATTGTGCATCAGGACA 240
Db 151 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATTGTGCATCAGGACA 92
QY 241 AGGGTA 246
Db 91 AGGGTA 86

RESULT 12

BG081002
LOCUS H3059H12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
DEFINITION H3059H12 5', mRNA sequence.
ACCESSION BG081002
VERSION BG081002
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 598)
REFERENCE Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,

Wood, W. H. III, Becker, K. G. and Ko, M. S. H.
 Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
 Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
 10922068
 On Jan 26, 2001 this sequence version replaced gi:12563570.
 Other ESTs: H3059H12-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@gsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3059 row: H column: 12
 Seq primer: -21M13 Reverse
 High quality sequence stop: 598
 POLYA=No.

FEATURES

source
 1. .598
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:H3059H12-5"
 /db_xref="taxon:10090"
 /clone="H3059H12"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 51.5%; Score 236.2; DB 2; Length 598;
 Best Local Similarity 71.3%; Pred. No. 5.9e-57;
 Matches 326; Conservative 0; Mismatches 128; Indels 3; Gaps 1;
 QY 1 ATGGCGTTAGAGTCTTGATGCTCCCTCGTGTCTTGATTTGGACCGTGCTCAGAACCTC 60
 Db 41 ATGGCTTTAGAGGCTTGTGGTTATCTTGTCTGCTTGTGGTTGACCTGCTGGGACATT 100
 QY 61 CATGTGAAATAAGTTGCTCTCGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
 Db 101 GAGCTGGATGAATGTTGTTCTCAGGACTGATGACGTCAGTGTAGTCGGTTTTCCAG 160
 QY 121 AGC---AGAAATCTGATATATTTCGGGATGAATTACATCTCGGAATGGCTGCCCTGCA 177
 Db 161 ACAAAGAAATCCATATATTTTGTGTGATGAATTATATCTGGGACAGAACTGCCGTGTA 220
 QY 178 AATCGATACATACATATGATATGATGTTATATATCTTGTTCGTGATTTGGCATCAGG 237
 Db 221 ACTCAATACAGCATCAATATGATTTTCATATACCTCTTCTCAGTGTGGCATCAGG 280
 QY 238 ACAAGGGTAGTTCTTGAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTTACCCCAAGG 297

Db 281 ACAAGGTTATCTCAATGAATCGTCTGTTTGAACCGAGATGATCTTCGTCCAAGG 340
 QY 298 AATATAGATCATGACCTCCAGAAATCATTTGGAGTGTTCACCTCTAGGAAATCAGTG 357
 Db 341 AATATATTGTTGGAACTTCAGATAGTCCTTTGCAGTGTTCTGCTCTCGGAAGTCTGTG 400
 QY 358 TGGCTTACACCACTTTCTACTGAGATGAATAAATTTGGATCTCTAGTCTCTTTTATGCT 417
 Db 401 TGGCTTATGCCACTATCTACTGAGGAGATCCAAACCTGTTAAGAGTCCCTTTATGACT 460
 QY 418 GACTTTAGACACACGACGAGAGAGTTAGGATATTATAA 454
 Db 461 GACTTTGAAGCTACACCAAGAGAGTTGGGTTATTATAA 497

RESULT 13

BY731676 629 bp mRNA linear EST 17-DEC-2002
 LOCUS
 DEFINITION
 ovary Mus musculus cDNA clone E330031A07 5', mRNA sequence.
 BY731676
 BY731676.1 GI:27144803
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 629)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Cousins, S., Dalla, E., Dragani, T.A.,
 Chotha, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gough, J., Grimmond, S.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawachi, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
 Kurochin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazawa, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, T., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851
 CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9226
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
 Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
 Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

TITLE

JOURNAL PUBMED COMMENT

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome* 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source

```

details.
Location/Qualifiers
1. .629
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E330031A07"
/sex="female"
/tissue_type="ovary"
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 2 days pregnant
adult female ovary"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5']
GAGAGAGAGCGGGCGGCACATCGAGTGTTCCTTTTTCCTTTTTCVN 3']. cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adaptor of sequence [5']
GAGAGAGAGATTCTGAGTATTAATTAATATCCCCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

```

ORIGIN

	Query Match	51.5%;	Score 236.2;	DB 5;	Length 629;
	Best Local Similarity	71.3%;	Pred. No. 6e-57;		
	Matches 326;	Conservative 0;	Mismatches 128;	Indels 3;	Gaps 1;
QY	1	ATGCGGTTAGAACTCTTGATGCTCCTCGCTGCTCTGATTTGGACCGGTCTGAGAACCTC	60		
DB	46	ATGCGCTTAAAGTCTCTGTTTATCTTCTGCTCTTGGTTTGGACTTGTGCTTGGGACATT	105		
QY	61	CATGTCAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA	120		
DB	106	GACCTGGATGAAGTTGTTCTCAGGACTGATCAGCGTCAGTCTTGTAGTCGTTTCCAG	165		
QY	121	AGC---AGAAATCTGTATATATTTTCGGAGTAATTCATCTGGGAATGGCTGCCCTGCA	177		
DB	166	AACAAAGAAATCCATATATTTTTGCTGATGAATATATCTGGGACAGAACTGCCGTGTA	225		
QY	178	AATCGGATACATATGTATATGATTTATATATCTTGTTCGTGATTGTGGCATCAGG	237		
DB	226	ACTCAGATACAGCACATCAATATGATTTCAATATACCTGTTTCTCACTGTGGCATCAGG	285		
QY	238	ACAAAGGTAGTTTCTGAGGAAACTCTCCCTTTTTTCAAAACCGAGCTGTACTTTATCCCAAGG	297		

	Db	286	ACAAAGTTTATCTCAAAATGAAATCGTCTGTTTGGAAACCGAGATGTACTTTCTGTC	CAAAGG	345
Qy <td>298<td>AATATAGATCATGACCCCTCAGGAAATCCATTTTGGAGTGTTCACACCTCTAGGAAATCAGTG<td>357</td><td></td><td></td></td></td>	298 <td>AATATAGATCATGACCCCTCAGGAAATCCATTTTGGAGTGTTCACACCTCTAGGAAATCAGTG<td>357</td><td></td><td></td></td>	AATATAGATCATGACCCCTCAGGAAATCCATTTTGGAGTGTTCACACCTCTAGGAAATCAGTG <td>357</td> <td></td> <td></td>	357		
Db <td>346<td>AATTTATGTTTGGAACTTTCAGATAGTCCCTTTGCAGTGTCTG<td>CCCTCTCGGAAGTCTGTG</td><td>405</td><td></td></td></td>	346 <td>AATTTATGTTTGGAACTTTCAGATAGTCCCTTTGCAGTGTCTG<td>CCCTCTCGGAAGTCTGTG</td><td>405</td><td></td></td>	AATTTATGTTTGGAACTTTCAGATAGTCCCTTTGCAGTGTCTG <td>CCCTCTCGGAAGTCTGTG</td> <td>405</td> <td></td>	CCCTCTCGGAAGTCTGTG	405	
Qy <td>358<td>TGGCTTTACACGATTTCTACTGAGATGAAATAAATTTGGATCCTTAGTCCTTTTATTTGCT</td><td>417</td><td></td><td></td></td>	358 <td>TGGCTTTACACGATTTCTACTGAGATGAAATAAATTTGGATCCTTAGTCCTTTTATTTGCT</td> <td>417</td> <td></td> <td></td>	TGGCTTTACACGATTTCTACTGAGATGAAATAAATTTGGATCCTTAGTCCTTTTATTTGCT	417		
Db <td>406<td>TGCGTTATGCGCACTCTCTACTGAGGAGGATCCAAAACCTGTTTAAAGAGTCCCTTTTATGACT</td><td>465</td><td></td><td></td></td>	406 <td>TGCGTTATGCGCACTCTCTACTGAGGAGGATCCAAAACCTGTTTAAAGAGTCCCTTTTATGACT</td> <td>465</td> <td></td> <td></td>	TGCGTTATGCGCACTCTCTACTGAGGAGGATCCAAAACCTGTTTAAAGAGTCCCTTTTATGACT	465		
Qy <td>418<td>GACTTTTCAGACAACAGCAGAGAGTTTAGGATTATTAA</td><td>454</td><td></td><td></td></td>	418 <td>GACTTTTCAGACAACAGCAGAGAGTTTAGGATTATTAA</td> <td>454</td> <td></td> <td></td>	GACTTTTCAGACAACAGCAGAGAGTTTAGGATTATTAA	454		
Db <td>466<td>GACTTTGAAGCTACACCAAGAGATTTGGGGTTATTAA</td><td>502</td><td></td><td></td></td>	466 <td>GACTTTGAAGCTACACCAAGAGATTTGGGGTTATTAA</td> <td>502</td> <td></td> <td></td>	GACTTTGAAGCTACACCAAGAGATTTGGGGTTATTAA	502		
RESULT 14					
CO810251					
LOCUS					
DEFINITION					
CO810251 862 bp mRNA linear EST 06-AUG-2004					
AGENCOURT_30259260 NIH_MGC_257 Mus musculus cDNA clone					
IMAGE:30931566 5', mRNA sequence.					

RESULT 14	ACCESSION	ORGANISM
CO810251	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FEATURES
source

ORIGIN

Query Match	51.1%;	Score 234.6;	DB 7;	Length 862;
Best Local Similarity	71.1%;	Pred. No. 1.9e-56;		
Matches 325;	Conservative 0;	Mismatches 129;	Indels 3;	Gaps 1;

QY	1	ATGGCGTTAAGAGTCTTGATGCTCTCGCTGCTTTGATTTGGACCGGTGCTGAGAACCTC	60
Db	8	ATGGCGTTTAGAGGTCCTGGTTTTTCTTGCTGCTTTGGTTTGACTTTGCTTTGGGACATT	67
QY	61	CATGTGAAATAAGTTTGCTCTCTGACACTGGTTTGATGGTCTCAGTTTATCCAGTTTCAGAA	120

D _b	68	GACGTGGATGTAAGTTGTTCTCAGACTGGAGA CGGCAGTGTTAGTCGGTTTTCC CAG	127
Q _y	121	AGC---AGAAATCTGTATATATTTCGCCGATGAATACTCGGAATGGGTGCCCTGCA	177
D _b	128	AACAAAAGAAATCCATATATTTTTTGCTGATGAATATATATCTGGGACAGAAC TGCCCGTGTA	187
Q _y	178	AATCGGATACACAATATATATATAGTATTATATATCTTGTTCGATGATGTGGCATCAGG	237
D _b	188	ACTCAAATACACGCACATCAATATGATTTCATACCCTGTTCTCACTGTGGCATCAGG	247
Q _y	238	ACAAGGGTAGTTTCTGAGGAAACCTCCCTTTTCAAACCGAGCTGACTTTACCCCAAGG	297
D _b	248	ACHAAGGTTATCTCAATAGAAATCGTCTGTGTTTTGNAACCGAGATGACTTTTCGTCCAAGG	307
Q _y	298	AATATAGATCATGACCCTCAGGAATAATCAATTTGGAGTGTTCCACTTAGGAAATCAGTG	357
D _b	308	AAITATTGTTTGGAACTTCAGATAGTCCTTATGTCAGTGTTCCTGCTCTCGGAAGTCTGTG	367
Q _y	358	TGGCTTACACCAAGTTTCTACTGAGATGAATAAATTTGGATCCTTAGTCCTTTTATTGCT	417
D _b	368	TGGCTTATGCCACTATCTACTTGAGAGGATCCAAAA CCTGTTAAGAGTCCCCTTTATGACT	427
Q _y	418	GACTTTTCAGACAACAGCAGAGAGTTAGGATTATTAA	454
D _b	428	GACTTTTGAAGCTCACACAGAGAGTTGGGGTTATTAA	464

RESULT	15
BB557840	
LOCUS	
DEFINITION	BB557840 RIKEN full-length enriched, 642 bp mRNA linear EST 26-OCT-2001 ovary Mus musculus cDNA clone E330031A07.3, mRNA sequence.
ACCESSION	BB557840
VERSION	BB557840.2 GI:16448937
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muriidae; Murinae; Mus.
REFERENCE	1 (bases 1 to 642) Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito.M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y. RIKEN Mouse ESTs (Arakawa.T., et al. 2001)
AUTHORS	TITLE

ORIGIN

Query Match	47.7%	Score 219;	DB 2;	Length 642;
Best Local Similarity	70.0%;	Pred. No. 5.7e-52;		
Matches	310;	Conservative 0;	Mismatches 130;	Indels 3; Gaps 1
QY	15	CTTGTACTCCTCGCTGCTCTGTGATTTTGGACCGGTGCTGAGAACCTCCATGTGAAATAAG	74	
Db	31	CTGGTTTATCTGCTGCTCTGGTTGGACATTTGTGCTTGGACATTTGACGTGGATGTTAG	90	
QY	75	TTGCTCTCTCGACTGGTGTGATGGTCTCAGTTTATCCAGTTTCGAGAAAGC--AGAAATCT	131	
Db	91	TTGTTCTCAGGACUGGATGACGGTCAGTTGTAGTGGCTTTTCCGAGACAAAGAATCC	150	
QY	132	GTATATATTGCGGATGAAATTACATCTGGGAATGGCTGCCCTGCAATTCGGATACATAC	191	
Db	151	ATATATTTTTCGCTGATGAAATTATATCTCGGACAGAACTGCCGTGTAACTCAGATACAGC	210	
QY	192	ATATGTATATGAGTTTATATATCTTTGTCGTGATTTGTGGCATCAGGACAAAGGTAGTTTC	251	
Db	211	ACATCAATATGATTTATATACCTGTTTCTACCTGTGGCATCAGACAAAGGTTATCTTC	270	
QY	252	TGAGGAAACTCTCCTTTTTCAAAACCGAGCTGACTTTTACCCCAAGGAATATAGATCATGA	311	
Db	271	AAATGAATCGTCGTGTTTGGAAACCGAGATGACTTTTCGTCCAAGGAATATTGTTTGA	330	
QY	312	CCCTCAGGAAATCCATTTGGAGTGTTCCA	CCTCTAGGAAATCAGTGTGGCTTACACCAGT	371
Db	331	ACTTCAGATAGTCCCTTTGACGTGTTCTGCCTCTCGAAGTCTGTGTGGCTTATGCCACT	390	

Qy	372	TTCTACTGAGATGAATATAAATTGGATCCTTAGTCCTTTTATTGCTGACTTTTCAGACAAC	431
Db	391	ATCTACTGAGGAGGATCCAAACCTGTTAAGAGTCCCTTTATGACTGACTTTGAAGCTAC	450
Qy	432	AGCAGAGAGCTTAGGATTATTAA	454
Db	451	ACCAGAAAGATTGGGGTTATTAA	473

Search completed: May 9, 2006, 14:40:04
Job time : 5673 secs

This Page Blank (uspto)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 9, 2006, 10:01:14 ; Search time 18.18 Seconds

(without alignments)

3327.968 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511

Perfect score: 826

Sequence: 1 atggcggttagaagctcttgat.....agttaggattattaatcttc 459

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-Q=/abs/ABSWEB/spool/US10664025/runat_08052006_173448_28291/app_query.fasta.1
-DB=A_Geneseq -QFMT=faстан -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs07
-USER=US10664025 @CGN_1_1_241 @runat_08052006_173448_28291 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	789	95.5	158	4	AAG89140 Human sec
2	789	95.5	158	5	ABG97353 Human CGD
3	789	95.5	158	5	ABG31324 Human GSS
4	789	95.5	158	6	ABR47931 Human sec
5	789	95.5	158	6	ABR00175 Human gen
6	789	95.5	158	7	ADB91675 Human sec
7	789	95.5	158	7	ADC74338 Human sec
8	789	95.5	159	3	ADB38337 Human sec
9	785	95.0	158	7	ADB65520 Human pro

10	775	93.8	148	6	ABP76238	Abp76238 Human GEN
11	752	91.0	143	6	ABP76237	Abp76237 Human GEN
12	429	51.9	106	4	AAG89145	Aag89145 Human sec
13	225.5	27.3	212	7	AD28663	Ad28663 Human NOV
14	225.5	27.3	212	8	ADM93408	Adm93408 Human NOV
15	223.5	27.1	212	2	Aaw73630	Aaw73630 Human sec
16	223.5	27.1	212	3	AAV94302	AAV94302 Human cor
17	223.5	27.1	212	4	AAW93517	AAW93517 Human pol
18	223.5	27.1	212	5	ABF62038	ABF62038 Human sec
19	223.5	27.1	212	5	ABF61861	ABF61861 Human pol
20	223.5	27.1	212	7	ADC24649	Adc24649 Protein e
21	223.5	27.1	212	8	ADL31210	Adl31210 Human pro
22	223.5	27.1	213	2	AAV12939	AAV12939 Amino aci
23	183.5	22.2	139	7	ABR83176	ABR83176 PLAC1 int
24	161	19.5	163	7	AD28665	Ad28665 Human NOV
25	161	19.5	163	8	ADM93410	Adm93410 Human NOV
26	98.5	11.9	930	4	ABG12048	Abg12048 Novel hum
27	96.5	11.7	139	7	ABR83177	ABR83177 PLAC1 int
28	94	11.4	59	4	ABR21608	ABR21608 Protein #
29	94	11.4	59	4	ABG51070	ABG51070 Human liv
30	88.5	10.7	234	4	ABG27342	Abg27342 Novel hum
31	85.5	10.4	1062	8	ADR09311	Adr09311 Human pro
32	85.5	10.4	3648	8	ADP30076	Adp30076 Human sec
33	85.5	10.4	3776	8	ADP30075	Adp30075 Human sec
34	85.5	10.4	5373	4	AAU14603	AAU14603 Novel bon
35	85.5	10.4	5373	7	ADJ68935	Adj68935 Human hea
36	85.5	10.4	5430	9	AD285102	Ad285102 Full-leng
37	85.5	10.4	5447	4	AAU14697	AAU14697 Novel bon
38	84	10.2	1241	6	ABU18092	Abu18092 Protein e
39	82.5	10.0	1122	2	AA64927	Aar64927 Cytadhesi
40	79.5	9.6	431	4	AAG85042	Aag85042 Shrimp wh
41	78.5	9.5	500	9	ADY22007	Ady22007 Antibody
42	77	9.3	358	5	ABB82478	Abb82478 Rat steat
43	77	9.3	358	7	AD561071	Ad561071 Rat Prote
44	77	9.3	358	8	ADI79900	Adi79900 Human E1-
45	76	9.2	158	8	ADK46733	Adk46733 Streptoco

ALIGNMENTS

RESULT 1

AAG89140

ID AAG89140 standard; protein; 158 AA.

AC AAG89140;

DT 11-SEP-2001 (first entry)

DE Human secreted protein, SEQ ID NO: 260.

KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis; GENSET.

OS Homo sapiens.

PN WO200142451-A2.

PD 14-JUN-2001.

PF 07-DEC-2000; 2000WO-IB001938.

PR 08-DEC-1999; 99US-0169629P.

PR 06-MAR-2000; 2000US-0187470P.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX WPI; 2001-367870/38.

XX N-PSDB; AAF64743.

PT Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of

diseases, and for diagnosis of those diseases.

Claim 21; Page 804; 921pp; English.

The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patient's own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention.

AA
SQ Sequence 158 AA;

Alignment Scores:		
Pred. No.:	8,95e-89	Length: 158
Score:	789.00	Matches: 151
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	95.5%	Indels: 0
DB:	4	Gaps: 0

US-10-664-025-43 COPY 53 511 (1-459) x AAG89140 (1-158)

Qy	1	ATGGCGTTAAGAGTCTTGATGCTCTCCGCTGCTTGATTTGGACCGGTGCTGAGAACCTC	60
Db	1	MetAlaLeuGluValIleuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu	20
Qy	61	CATGTCAAAATAAGTTGCTCTCTGACATGGTTGATGGTCTCAGTATATCCAGTTGCAGAA	120
Db	21	HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
Qy	121	AGCAGAAATCTGTATATATTTGCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAAAT	180
Db	41	SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn	60
Qy	181	CGGATACATACATATGATATCAGTGTATATATATCTTGTTGCTGATTCTGGGCATCAGGACA	240
Db	61	ArgIleHisThrTyrlValTyrlGluPheIleTyrlLeuValArgAspCysGlyIleArgThr	80
Qy	241	AGSGTAGTTTCTGAGGAAACTCTCCTTTTTCAAACCGAGCTGTATCTTTACCCCAAGGAAT	300
Db	81	ArgValIleSerGluGluThrLeuLeuPheGlnThrGluLeuTyrlPheThrProArgAsn	100
Qy	301	ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCCACCTCTAGGAAATCAGTGTGG	360
Db	101	IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp	120
Qy	361	CTTACACCAAGTTTCTACTGAGAAATCAAAATAAAATTTGGATCCTAGTCCCTTTATTGTGTGAC	420
Db	121	LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp	140
Qy	421	TTTTCAGACAACAGCAGAGAGTAGGATATTATTA	453
Db	141	PheGlnThrThrAlaGluLeuGlyLeuLeu	151

RESULT 2

ABG97353

ID	ABG97353 standard; protein; 158 AA.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

AC ABG97353;

XX


```
CC sequences of proteins associated with CGDD. The present sequence
XX represents a CGDD protein
SQ Sequence 158 AA;

Alignment Scores:
Pred. No.:      8,95e-89      Length:      158
Score:          789.00      Matches:      151
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      95.5%      Indels:      0
DB:              5      Gaps:      0

US-10-664-025-43_COPY_53_511 (1-459) x ABG97353 (1-158)
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAAsnLeu 20
QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGCTGATGGTCTCAGTTATCCAGTTGCAGAA 120
DB 21 HisValIysIleSerCysSerLeuAAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTGGCGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 41 SerArgAsnLeuTyrllePheAlaAAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATGATTATATATCTTGTTCGTTGATTTGGCATCAGGACA 240
DB 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAAspCysGlyIleAsgThr 80
QY 241 AGGGTAGTTCTCAGGAAACTCTCCTTTTCAACCGAGCTGTACTTACCCCAAGGAAT 300
DB 81 ArgValIleSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAATCAGGTGG 360
DB 101 IleAspHisAAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTrp 120
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAATAAATTCGATCTAGTCTTATTTATGCTGAC 420
DB 121 LeuThrProValSerThrGluAAsnGluIleLysLeuAAspProSerProPheIleAlaAAsp 140
QY 421 TTTACAGACACACAGAGAGTTAGGATTATTA 453
DB 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 3
ABG31324
ID ABG31324 standard; protein; 158 AA.
XX AC ABG31324;
XX DT 05-NOV-2002 (first entry)
XX DE Human GSSP3 polypeptide.
XX KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
KW body mass; serum glucose regulation; body weight loss; obesity;
KW metabolic-related disorder; impaired glucose tolerance; stroke;
KW insulin resistance; hyperlipidemia; atherosclerosis; heart disease;
KW hypertension; syndrome C; type I diabetes; type II diabetes;
KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
KW polycystic ovarian syndrome; acanthosis nigrican; leprechaunism;
KW lipatrophy; physical performance; exercise; dyslexia; schizophrenia;
KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT 1. .17
XX FT /label= Signal_peptide
XX FT 18. .158
```

```
FT
XX
XX
XX WO200260466-A2.
XX PD 08-AUG-2002.
XX PF 01-FEB-2002; 2002WO-IB001333.
XX PR 02-FEB-2001; 2001US-0266156P.
XX PA (GEST ) GENSET.
XX XX
XX Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
XX BiHain B;
XX WPI; 2002-608487/65.
XX DR N-PSDB; ABK90053.
XX PT Reducing circulating glucose levels or increasing insulin sensitivity,
XX useful for reducing body mass or preventing body weight gain, comprises
XX administering composition comprising GSSP3 polypeptide.
XX
XX Claim 1; Page 96-97; 97pp; English.
XX
XX The present invention relates to the isolation of human GSSP3
XX polypeptide, and polynucleotide sequences that encode it. The GSSP3
XX polypeptide reduces circulating blood glucose levels, increases insulin
XX sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
XX polynucleotide sequences are useful in serum glucose regulation and
XX acid metabolism, body weight loss, and prevention of body weight gain.
XX Compositions comprising GSSP3 polypeptides are useful for controlling
XX blood glucose levels, for treating metabolic-related diseases or
XX disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
XX hyperlipidemia, atherosclerosis, heart disease, hypertension, stroke,
XX syndrome C, type I or II diabetes, diabetes related complications,
XX microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
XX syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
XX acanthosis nigricans, leprechaunism and lipostrophy). The polypeptides
XX are also useful to improve physical performance during work or exercise,
XX and to treat dyslexia, attention-deficit disorder, attention-
XX deficit/hyperactivity disorder, and psychiatric disorders such as
XX schizophrenia. The present sequence represents human GSSP3 polypeptide
XX
XX Sequence 158 AA;

Alignment Scores:
Pred. No.:      8,95e-89      Length:      158
Score:          789.00      Matches:      151
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      95.5%      Indels:      0
DB:              5      Gaps:      0

US-10-664-025-43_COPY_53_511 (1-459) x ABG31324 (1-158)
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAAsnLeu 20
QY 61 CATGTGAAAATAAGTTGCTCTCTGGACTGCTGATGGTCTCAGTTATCCAGTTGCAGAA 120
DB 21 HisValIysIleSerCysSerLeuAAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCGCGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 180
DB 41 SerArgAsnLeuTyrllePheAlaAAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATGATTATATATCTTGTTCGTTGATTTGGCATCAGGACA 240
DB 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAAspCysGlyIleAsgThr 80
QY 241 AGGGTAGTTCTCAGGAAACTCTCCTTTTCAACCGAGCTGTACTTACCCCAAGGAAT 300
DB 81 ArgValIleSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAATCAGGTGG 360
DB 101 IleAspHisAAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTrp 120
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAATAAATTCGATCTAGTCTTATTTATGCTGAC 420
DB 121 LeuThrProValSerThrGluAAsnGluIleLysLeuAAspProSerProPheIleAlaAAsp 140
QY 421 TTTACAGACACACAGAGAGTTAGGATTATTA 453
DB 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
```

Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
QY 361 CTTACACCATGTTCTTACTGAGATGAAATGAAATTTGGATCCCTAGTCTTTTATTCGTGAC 420
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTCAGACACAGCAGAGAGTATGATTATTA 453
Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 4
ABR47931
ID ABR47931 standard; protein; 158 AA.
XX
AC ABR47931;
DT 12-JUN-2003 (first entry)
DE Human secreted protein, SEQ ID 822.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
KW vulnery; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX
OS Homo sapiens.
XX
PN WO200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002WO-US009785.
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI WPI; 2003-129429/12.
DR
XX Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
PT disorders such as arrhythmia.
XX
PS Claim 13; SEQ ID NO 822; 1881pp; English.
XX

The present invention relates to novel human secreted proteins (ABR47633-
CC ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
CC and their coding sequences are useful for the preparation of a diagnostic
CC or pharmaceutical composition for diagnosing or treating a cardiovascular
CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
CC arteriosclerosis and myocardial ischaemia), neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in

CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 158 AA;
Alignment Scores:
Pred. No.: 8,95e-89 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: Gaps: 0
US-10-664-025-43_copy_53_511 (1-459) x ABR47931 (1-158)
QY 1 ATGGCGTTAGAGTCTTCTGATGCTCGCTGCTTGTGATTTGGACCGGTGCTGAGACCTC 60
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAAATAAGTTGCTCTCTGACCTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 21 HisValLysLleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCGGATGAAATACATCTGGGAATGGCTGCCCTGCAAT 180
Db 41 SerArgAsnLeuTyrrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATGATTTATATATCTTGTTCGTGATTTGGCATCAGGACA 240
Db 61 ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGCTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 300
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
QY 361 CTTACACCATGTTCTTACTGAGATGAAATGAAATTTGGATCCCTAGTCTTTTATTCGTGAC 420
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTCAGACACAGCAGAGAGTATGATTATTA 453
Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 5
ABR00175
ID ABR00175 standard; protein; 158 AA.
XX
XX ABR00175;
AC ABR00175;
XX
DT 03-APR-2003 (first entry)
XX
DE Human gene 165 encoded secreted protein HTLS08, SEQ ID NO:464.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
XX WO200276488-A1.
XX
XX 03-OCT-2002.
XX
XX 19-MAR-2002; 2002WO-US008276.
XX
XX 21-MAR-2001; 2001US-0277340P.
PR

PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX WPI; 2003-029900/02.
DR N-PSDB; ABZ71354.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers.
XX
XX Claim 13; Page 1048; 1216pp; English.
XX
CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein of the
CC invention
XX
SQ Sequence 158 AA;

Alignment Scores:
Pred. No.: 8,95e-89 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x ABR00175 (1-158)

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAenLeu 20
QY 61 CATGTGAAATAAGTTGCTCTGGACTGTTGATGCTCTCAGTTATCCAGTTGCGAGAA 120
DB 21 HisValLysIleSerCysSerLeuAenTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTGCGGATGAATACATCTCGGAATGGGCTGCCCTGCAAA 180
DB 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAen 60
QY 181 CGGATACATACATATGATATAGTTTATATATATCTTGTGCTGATTTGGCATFCAGACA 240
DB 61 ArgIleHisThrTyrlleValTyrllePheIleTyrlleValArgAspCysGlyIleArgThr 80
QY 241 AGGGTAGTTCTTGAGAAACTCTCTTTTTCACACCGAGCTGTACTTTACCCNAGGAAT 300
DB 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAen 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120

QY 361 CTTACACCAAGTTTCTACTGGAATGAATAAATTCGATCCTAGTCTCTTTATTGCTGAC 420
DB 121 LeuThrProValSerThrGluAenGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTTCAGACACAGCAGAGAGAGTTAGGATTATTA 453
DB 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 6
ADB91675
ID ADB91675 standard; protein; 158 AA.
AC ADB91675;
XX
XX 04-DEC-2003 (first entry)
XX Human secreted protein #SEQ ID 621.
XX Secreted protein; gene therapy; antidiabetic; diabetes; human.
XX Homo sapiens.
XX WO2003004622-A2.
XX 16-JAN-2003.
XX 19-MAR-2002; 2002WO-US008124.
XX 21-MAR-2001; 2001US-0277340P.
XX 19-JUL-2001; 2001US-0306171P.
XX 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2003-229407/22.
XX
XX Nucleic acid encoding a human secreted protein is useful in diagnosing or
XX treating diabetes or conditions related to diabetes.
XX Claim 3; SEQ ID NO 621; 1537pp; English.
XX The invention relates to isolated nucleic acid molecules ADB91065-
XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX ADB91834. Also disclosed is a recombinant vector comprising a
XX polynucleotide of the invention, and a recombinant host cell comprising
XX the recombinant vector. The polypeptide of the invention is useful in
XX identifying a binding partner by contacting the polypeptide with a
XX binding partner, and determining whether the binding partner increases or
XX decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX antibody or its fragment, agonist or antagonist are useful for preparing
XX a pharmaceutical composition for diagnosing or treating diabetes or
XX conditions related to diabetes. The present sequence is that of the human
XX immunoglobulin Fc portion used to generate fusion proteins, increasing
XX the stability of the fused protein as compared to the secreted protein
XX only. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 158 AA;

Alignment Scores:
Pred. No.: 8,95e-89 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x ADB91675 (1-158)

QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTCTGAGAACCTC 60

```
Db      |||||MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuIleTrpThrGlyAlaGluAsnLeu 20
QY      61 CATGTGAAATAAGTTGCTCTCTGACCTGGTTGATGGTCTCAGTTATCCAGTTCAGAA 120
Db      21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY      121 AGCAGAAATCTGTATATATTTCCGGATGAATTACATCTGGGAATGGGCTGCCCTCAAA 180
Db      41 SerArgAsnLeuTyriIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY      181 CGGATACATACATATGATATATAGATTATATATCTTTGTCGTGATGTCGATGCGCATCAG 240
Db      61 ArgIleHisThrTyriValTyriGluPheIleTyriLeuValArgAspCysGlyIleArgThr 80
QY      241 AGGGTAGTTTCTGAGGAACTCTCTTTTCAACCGAGCTGCTACTTACCCCAAGGAAT 300
Db      81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyriPheThrProArgAsn 100
QY      301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
QY      361 CTTACACAGTTTCTTACTGAGAAATGAATAATAATTGGATCTCTAGTCTTTTATCTGCTGAC 420
Db      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY      421 TTTTCAGACAAACAGCAGAGAGTTAGGATTATTA 453
Db      141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 7
ADC74338
ID ADC74338 standard; protein; 158 AA.
XX
AC ADC74338;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human secreted protein - SEQ ID 971.
XX
KW antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
KW antidiabetic; immunosuppressive; dermatological; nephrotropic;
KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;
KW haemopoietic; haematologic; anaemia; autoimmune disorder;
KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
KW human.
XX
OS Homo sapiens.
XX
XX WO2003038063-A2.
XX
XX 08-MAY-2003.
XX
XX 19-MAR-2002; 2002WO-US008277.
XX
XX 21-MAR-2001; 2001US-0277340P.
XX
XX 19-JUL-2001; 2001US-0306171P.
XX
XX 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-430516/40.
XX
XX N-PSDB; ADC73723.
XX
XX New human secreted polypeptide for diagnosing, preventing or treating
XX hematopoietic or hematologic disorders (e.g. anemia), autoimmune
PT
```

```
PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
PT atherosclerosis).
XX
PS Claim 16; SEQ ID NO 971; 2272pp; English.
XX
CC The invention relates to a novel human secreted polypeptide comprising a
CC defined sequence given in the specification. The polypeptide, nucleic
CC acid molecule, antibody, agonist or antagonist of the invention may be
CC useful for preparing a composition for diagnosing or treating a
CC haemopoietic or haematologic disorder such as anaemia, autoimmune
CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
CC diabetes, systemic lupus erythematosus or glomerulonephritis,
CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
CC disease, wounds and hyperproliferative disorders including
CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
CC parasitic infections. The polypeptide may also be used during gene
CC therapy procedures and for identifying a binding partner by contacting
CC the polypeptide with a binding partner and determining whether the
CC binding partner increases or decreases the activity of the polypeptide.
CC The current sequence is that of the human secreted protein of the
CC invention.
XX
SQ Sequence 158 AA;

Alignment Scores:
Pred. No.:      8.95e-89      Length:      158
Score:          789.00      Matches:      151
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      95.5%      Indels:      0
DB:              7      Gaps:      0

US-10-664-025-43_COPY_53_511 (1-459) x ADC74338 (1-158)

QY      1 ATGGCGTTAGAGTCTTCATGCTCTCGCTGCTTGTGATTGGACCGGTGCTGAGAACCTC 60
Db      1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY      61 CATGTGAAATAAGTTGCTCTCTGACCTGGTTGATGGTCTCAGTTATCCAGTTCAGAA 120
Db      21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY      121 AGCAGAAATCTGTATATATTTCCGGATGAATTACATCTGGGAATGGGCTGCCCTCAAA 180
Db      41 SerArgAsnLeuTyriIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY      181 CGGATACATACATATGATATATAGATTATATATCTTTGTCGTGATGTCGATGCGCATCAG 240
Db      61 ArgIleHisThrTyriValTyriGluPheIleTyriLeuValArgAspCysGlyIleArgThr 80
QY      241 AGGGTAGTTTCTGAGGAACTCTCTTTTCAACCGAGCTGCTACTTACCCCAAGGAAT 300
Db      81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyriPheThrProArgAsn 100
QY      301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
QY      361 CTTACACAGTTTCTTACTGAGAAATGAATAATAATTGGATCTCTAGTCTTTTATCTGCTGAC 420
Db      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY      421 TTTTCAGACAAACAGCAGAGAGTTAGGATTATTA 453
Db      141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 8
AAB38337
ID AAB38337 standard; protein; 159 AA.
XX
XX AAB38337;
XX
XX 31-JAN-2001 (first entry)
DT
```

XX DE Human secreted protein encoded by gene 17 clone HTELS08.

XX KW Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;

XX KW cystostatic; cardiac; vasotropic; cerebroprotective; neuroprotective;

XX KW neutropic; antibacterial; virucide; fungicide; ophthalmological; human;

XX KW vulnery; gene therapy; infection; secreted protein.

XX OS Homo sapiens.

XX PN W0200061623-A1.

XX PD 19-OCT-2000.

XX PF 06-APR-2000; 2000WO-US008979.

XX PR 09-APR-1999; 99US-0128693P.

XX PR 26-APR-1999; 99US-0130991P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;

XX PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;

XX PI Young PE;

XX DR WPI; 2000-647418/62.

XX PT New nucleic acid molecules encoding 62 human secreted proteins for

XX PT diagnosing, preventing, treating or ameliorating medical conditions and

XX PT used as food additives or preservatives.

XX PS Claim 11; Page 608-609; 716pp; English.

XX CC Sequences AAB38321-B38396 represent the amino acid sequences of 62 human

XX CC secreted proteins encoded by the genes AAC69512-C69587. The genes and

XX CC proteins are useful for preventing, ameliorating or treating medical

XX CC conditions, e.g. by protein or gene therapy. The genes are isolated from

XX CC a range of human tissues disclosed in the specification. The nucleic

XX CC acids, proteins, antibodies and (ant)agonists are useful in the

XX CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.

XX CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of

XX CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;

XX CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angioneurosis;

XX CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections

XX CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.

XX CC corneal infection. The polypeptides can also be used to aid wound healing

XX CC and epithelial cell proliferation, to prevent skin aging due to sunburn,

XX CC to maintain organs before transplantation, for supporting cell culture of

XX CC primary tissues, to regenerate tissues and in chemotaxis

XX SQ Sequence 159 AA;

Alignment Scores:

Pred. No.: 8,98e-89 Length: 159

Score: 789.00 Matches: 151

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 95.5% Indels: 0

DB: 3 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x AAB38337 (1-159)

QY 1 ATGGCGTTTGAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTCTGAGAACCTC 60

Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20

QY 61 CATGTCAAAATAGTTGCTCTGGACTGGTTCAGTCTCAGTTATCCAGTTGCAGAA 120

Db 21 HsValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40

QY 121 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGCTGCCCTGCAAT 180

Db 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60

QY 181 CGGATACATACATATGTATATATGATTTATATATCTTGTCTGATTTGGCATCAGACA 240

Db 61 ArgIleHisThrTyValTyGluPheIleTyLeuValArgAspCysGlyIleArgThr 80

QY 241 AGGCTAGTTTCTCAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300

Db 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn 100

QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTTAGGAAATCAGTGTGG 360

Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTtp 120

QY 361 CTTACACCACTTTCTACTGAGATGAATAAATTTGATCCTAGTCTCTTTTATTCCTGAC 420

Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140

QY 421 TTTTCAGACCAACAGCAGAGAGATTAGGATTATTA 453

Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 9

ADB65520

ID ADB65520 standard; protein; 158 AA.

XX AC ADB65520;

XX XX

DT 04-DEC-2003 (first entry)

XX XX

DE Human protein encoded by clone TESTI20282420.

XX XX

KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;

KW cell regeneration; membrane protein; signal transduction-related protein;

KW transcription-related protein; osteoporosis; neurological disease;

KW cancer; tumour.

XX XX

OS Homo sapiens.

XX XX

PN EP1308459-A2.

XX XX

PD 07-MAY-2003.

XX XX

PF 28-MAR-2002; 2002EP-00007401.

XX XX

PR 05-NOV-2001; 2001JP-00379298.

XX XX

PR 25-JAN-2002; 2002US-00350978.

XX XX

PA (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX XX

DR WPI; 2003-450961/43.

DR N-PSDB; ADB63550.

XX XX

PT New polynucleotides and polypeptides, useful for developing a diagnostic

PT marker or medicines for regulation of their expression and activity, or

PT as targets of gene therapy.

XX XX

PS Claim 1; Page; 222pp; English.

XX XX

CC The invention discloses a polynucleotide comprising a sequence selected

CC from 1970 fully defined nucleotide sequences which encode novel

CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide

CC or its partial peptide, an antibody binding to the polypeptide or peptide

CC of the polynucleotide, immunologically assaying the polypeptide or

CC peptide of the polynucleotide by contacting the polypeptide or peptide

CC with the antibody of the encoded protein, and observing the binding

CC between the two, a transformant carrying the polynucleotide in an

CC expressible manner and an antisense polynucleotide. The oligonucleotide

CC is useful as a primer for synthesising the polynucleotide, or as a probe

CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 158 AA;

Alignment Scores:
Pred. No.: 2,82e-88 Length: 158
Score: 785.00 Matches: 150
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 95.0% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x ADB65520 (1-158)

QY 1 ATGGCGTTAGAGTCTTGTATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTCAGAA 120
DB 21 HisValIleSerCysSerLeuAspTrpLeuMetAlaSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
DB 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATATATATCTTGTTCGTGATGGGATCGGATCAGGACA 240
DB 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGCTAGTTTCTGAGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
DB 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
QY 301 APAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
QY 361 CTTACACAGTTTCTACTGAGAAATGAATAAAATTTGGATCCCTAGTCCCTTTTATTCCTGAC 420
DB 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTTCAGACACAGCAGAGAGTAGATTATTA 453
DB 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 10

ABP76238

ID ABP76238 standard; protein; 148 AA.

XX AC ABP76238;

XX

XX 21-FEB-2003 (first entry)

XX Human GENSET protein SEQ ID 788.

XX

XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;

XX gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;

XX inflammatory disease; immune disorder; neuromuscular; toxicity;

XX central nervous system; cardiovascular; gastrointestinal.

XX Homo sapiens.
OS WO200283898-A1.
XX
XX 24-OCT-2002.
XX
XX 18-APR-2001; 2001WO-IB000914.
XX
XX 18-APR-2001; 2001WO-IB000914.
XX (GEST) GENSET.
XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX WPI; 2003-075548/07.
XX
XX New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
PS Claim 14; Page 680; 735pp; English.
XX
CC The present invention relates to novel GENSET polynucleotides (AB236404-
CC AB236911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity
XX
SQ Sequence 148 AA;

Alignment Scores:
Pred. No.: 4,84e-87 Length: 148
Score: 775.00 Matches: 148
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 93.8% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x ABP76238 (1-148)

QY 1 ATGGCGTTAGAGTCTTGTATGCTCTCGCTCTCTTGTGATTTGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTCAGAA 120
DB 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
DB 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATATATATCTTGTTCGTGATGGGATCGGATCAGGACA 240
DB 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGCTAGTTTCTGAGAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 300
DB 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
QY 301 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
QY 361 CTTACACAGTTTCTACTGAGAAATGAATAAAATTTGGATCCCTAGTCCCTTTTATTCCTGAC 420
DB 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140

421 TTTCAGACACAGCAGAGATTA 444
141 PheGlnThrAlaGluGluLeu 148

RESULT 11
ABP76237
ID ABP76237 standard; protein; 143 AA.
XX AC ABP76237;
XX
XX 21-FEB-2003 (first entry)
XX
XX Human GENSET protein SEQ ID 787.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
inflammatory disease; immune disorder; neuromuscular; toxicity;
central nervous system; cardiovascular; gastrointestinal.
XX
OS Homo sapiens.
XX
XX WO200283898-A1.
XX
XX 24-OCT-2002.
XX
XX 18-APR-2001; 2001WO-IB000914.
XX
XX 18-APR-2001; 2001WO-IB000914.
XX
XX (GEST) GENSET.
XX
XX Benjamin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX WPI; 2003-075548/07.
XX
XX
XX New GENSET polynucleotides and polypeptides, useful for treating heavy
metal toxicity, cancer, inflammatory diseases, immune disorders, and the
neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
toxicity.
XX
XX
XX Claim 14; Page 680; 735pp; English.
XX
XX The present invention relates to novel GENSET polynucleotides (ABZ36404-
ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
diseases, immune disorders, and the neuromuscular, central nervous system
(CNS), cardiovascular or gastrointestinal effects of the toxicity
XX
XX Sequence 143 AA;

Alignment Scores:
Pred. No.: 3.5e-84 Length: 143
Score: 752.00 Matches: 143
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x ABP76237 (1-143)

QY 1 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTCTGAGACCTC 60
Db 1 MetAlaLeuGluValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 61 CATGTGAAAATAAGTTGGTCTCTCGACCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 21 HisValIleValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTCTATATATTTGCGGATGAATTAATCTGCGAATGGGCTGCCCTGCAAT 180

CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET polypeptide of
CC the invention
XX
SQ Sequence 106 AA;

Alignment Scores:
Pred. No.: 5,27e-44 Length: 106
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 51.9% Indels: 0
DB: 4 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x AAG89145 (1-106)

QY 1 ATGGCGTTAAGAGTCTTGATGCTCTCGCTGCTGTTGATTGACCGGTCTGAGAACCTTC 60
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTyrThrGlyAlaGluAenLeu 20
QY 61 CATGTGAAATTAAGTGTCTCTGACCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 120
Db 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTCGGAATGGGCTGCCCTGCAAT 180
Db 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGGTA 246
Db 81 ArgVal 82

RESULT 13
ADE28663
ID ADE28663 standard; protein; 212 AA.

XX AC ADE28663;
XX DT 29-JAN-2004 (first entry)
XX DE Human NOV12a protein - SEQ ID 40.
XX KW NOVX; antidiabetic; anorectic; cardiatic; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antiinflammatory; dermatological;
KW antiasthmatic; antipneumic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; NOV.

OS Homo sapiens.
XX
XX WO2003040330-A2.
XX
XX PD 15-MAY-2003.
XX
XX PF 05-NOV-2002; 2002WO-US035536.
XX
XX PR 05-NOV-2001; 2001US-0338626P.
PR 05-DEC-2001; 2001US-0336500P.
PR 07-DEC-2001; 2001US-0338285P.
PR 12-DEC-2001; 2001US-0341346P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.

PR 20-DEC-2001; 2001US-0342592P.
PR 27-DEC-2001; 2001US-0344297P.
PR 31-DEC-2001; 2001US-0344903P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0384024P.
PR 07-AUG-2002; 2002US-0401788P.
PR 26-AUG-2002; 2002US-0406353P.
PR 31-OCT-2002; 2002US-00287971.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boidog FL;
PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K;
PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI Pena CA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA;
PI Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VI, Twomlow N;
PI Vernet CAM, Zerhusen BD, Zhong M;
XX
XX WPI: 2003-441555/41.
DR N-PSDB; ADE28662.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
XX Claim 1: SEQ ID NO 40; 447pp; English.
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC cardiatic, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC dermatological, antiasthmatic and antipneumic activities. The
CC polypeptides, nucleic acid molecules and antibodies may be useful for
CC treating or diagnosing diseases including metabolic disorders such as
CC diabetes and obesity, infectious diseases, anorexia, cancer,
CC cardiovascular diseases including hypertension and atherosclerosis,
CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC Furthermore, the nucleic acids and polypeptides may also be used to
CC identify molecules that modulate or inhibit neurogenesis, cell
CC differentiation and proliferation, haemopoiesis, wound healing and
CC angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC be used as hybridisation probes, in chromosome mapping, tissue typing,
CC preventive medicine and pharmacogenomics. The current sequence is that of
CC the human NOV protein of the invention.
XX
XX Sequence 212 AA;

Alignment Scores:
Pred. No.: 1.52e-18 Length: 212
Score: 225.50 Matches: 40
Percent Similarity: 60.8% Conservative: 33
Best Local Similarity: 33.3% Mismatches: 42
Query Match: 27.3% Indels: 5
DB: 7 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x ADE28663 (1-212)

QY 13 GTCTTGATGCTCTCGCTGCTGTTGATTTGGACCGGTCTGAGAACCTCCATCGAAATA 72
Db 9 LeuMetIleLeuLeuThrSerAlaLeuSerAlaGlySerGlyClnSerProMetThrVal 28

CC polypeptides, polynucleotides and antibodies that bind immunospecifically
CC to the polypeptides are useful for treating or preventing disorders or
CC syndromes such as congenital heart defects, cardiomyopathy,
CC atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC anorexia and immune disorders. This sequence represents a human NOVX
CC polypeptide of the invention. Note: The sequence data for this patent is
CC also available from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX

SQ Sequence 212 AA;

Alignment Scores:
Pred. No.: 1,52e-18 Length: 212
Score: 225.50 Matches: 40
Percent Similarity: 60.8% Conservative: 33
Best Local Similarity: 33.3% Mismatches: 42
Query Match: 27.3% Indels: 5
DB: 8 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x ADM93408 (1-212)

QY 13 GTCCTGATGCTCTCGCTCTCTGTTGACCGGTGCTGAGAACCTCCATGTAATAA 72
DB 9 LeuMetIleLeuLeuThrSerAlaLeuSerAlaGlySerGlyGlnSerProMetThrVal 28
QY 73 AGTTGCTCTCGACTGGTGTGATGCTCAGTTATCCAGTTGCGAGAAAGCAGAAATCG 132
DB 29 LeuCysSerIleAspTrpPheMetValThrValHisProPheMetLeuAsnAspVal 48
QY 133 TATATATTTGCGGATCAATATACATCTGGAAATGGCTGCCCTGCCAATCGGATACATACA 192
DB 49 CysValHisPheHisGluLeuHisLeuGlyLeuGlyCysProProAsnHisValGlnPro 68
QY 193 TATGATATGAGTTTATATATCTTGTGTTGATTTGGCATCAGACAGGAGTTTCT 252
DB 69 HisAlaTyrGlnPheThrTyrArgValThrGluCysGlyIleArgAlaLysAlaValSer 88
QY 253 GAGGAAACTCTCTTTTCAACCCGAGCTGTACTTTTACCCCAAGGAATATAGATCATGAC 312
DB 89 GlnAspMetValIleTyrSerThrGluIleHisTyrSerSerLysGly-----Thr 105
QY 313 CCTCAGGAA-----ATCCATTGGAGTGTCTCCACTCTAGGAATCAGTGTGGCTTACA 366
DB 106 ProSerLysPheValIleProValSerCysAlaAlaProGlnLysSerProTrpLeuThr 125

RESULT 15

AAW73630

ID AAW73630 standard; protein; 212 AA.

AC AAW73630;

DT 23-MAR-1999 (first entry)

DE Human secreted protein clone ej265_4.

XX Secreted protein; human; nutritional supplements; immune stimulant;
KW immune suppressor; haematopoiesis regulator; tissue growth; chemotaxis;
KW activin/inhibin; chemokinesis; haemostasis; thrombolysis;
KW receptor/ligand activity; anti-inflammatory; tumour inhibitor;
KW cadherin/tumour invasion suppressor.

XX Homo sapiens.

OS

PN WO9855614-A2.

XX 10-DEC-1998.

XX 01-JUN-1998; 98WO-011210.

XX 04-JUN-1997; 97US-00868696.

PR 04-JUN-1997; 97US-00868697.
PR 04-JUN-1997; 97US-00868698.
PR 04-JUN-1997; 97US-00868698.
PR 04-JUN-1997; 97US-00868699.
PR 04-JUN-1997; 97US-00868699.
PR 04-JUN-1997; 97US-00869191.
PR 04-JUN-1997; 97US-00869191.
PR 04-JUN-1997; 97US-00869192.
PR 04-JUN-1997; 97US-00869193.
PR 04-JUN-1997; 97US-00869194.
PR 29-MAY-1998; 98US-0087252P.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCOY JM, Lavallie ER, Racie LA, Treacy M, Spaulding V;
PI Agostino MJ, Howes SH, Fechtel K;
XX N-PSDB; AAV55748.

DR WPI; 1999-059912/05.

XX N-PSDB; AAV55748.

PT New polynucleotides encoding secreted human proteins - derived from human
PT foetal brain, adult brain, foetal kidney, adult ovary, adult retina,
PT adult placenta or adult uterus cDNA libraries.

PS Claim 31; Page 93-94; 127pp; English.

XX This sequence represents a human secreted protein of the invention. The
CC DNA encoding this sequence was isolated from a human adult placenta cDNA
CC library, and was designated clone ej265_4. The DNAs and proteins are
CC predicted to have biological activities which would make them suitable
CC for treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional sources or supplements, immune stimulating or
CC suppressing activity, haematopoiesis regulating activity, tissue growth
CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
CC haemostatic and thrombolytic activity, receptor/ligand activity, anti-
CC inflammatory activity, cadherin/tumour invasion suppressor activity, and
CC tumour inhibition activity. The DNAs are also stated to be useful for
CC gene therapy. A host cell transfected with the DNA, or its subfragments
CC and variants is useful for recombinant production of the human secreted
CC protein clones

XX Sequence 212 AA;

SQ Alignment Scores:
Pred. No.: 2,69e-18 Length: 212
Score: 223.50 Matches: 40
Percent Similarity: 60.0% Conservative: 32
Best Local Similarity: 33.3% Mismatches: 43
Query Match: 27.1% Indels: 5
DB: 2 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x AAW73630 (1-212)

QY 13 GTCCTGATGCTCTCGCTCTCTGTTGACCGGTGCTGAGAACCTCCATGTAATAA 72

DB 9 LeuMetIleLeuLeuThrSerAlaPheSerAlaGlySerGlyGlnSerProMetThrVal 28

QY 73 AGTTGCTCTCGACTGGTGTGATGCTCAGTTATCCAGTTGCGAGAAAGCAGAAATCG 132

DB 29 LeuCysSerIleAspTrpPheMetValThrValHisProPheMetLeuAsnAspVal 48

QY 133 TATATATTTGCGGATCAATATACATCTGGGAATGGCTGCCCTGCCAATCGGATACATACA 192

DB 49 CysValHisPheHisGluLeuHisLeuGlyLeuGlyCysProProAsnHisValGlnPro 68

QY 193 TATGATATGAGTTTATATATCTTGTGTTGATTTGGCATCAGACAGGAGTTTCT 252

DB 69 HisAlaTyrGlnPheThrTyrArgValThrGluCysGlyIleArgAlaLysAlaValSer 88

QY 253 GAGGAAACTCTCTTTTCAACCCGAGCTGTACTTTTACCCCAAGGAATATAGATCATGAC 312

DB 89 GlnAspMetValIleTyrSerThrGluIleHisTyrSerSerLysGly-----Thr 105

Search completed: May 9, 2006, 10:04:48
Job time : 92.9 secs

This Page Blank (Intentional)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:08:38 ; Search time 7.02 Seconds
(without alignments)
1621.715 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511

Perfect score: 826

Sequence: 1 atggcgtagaagtctgtat.....agttaggattataattcttc 459

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DRV=xlp
-Q=/abs/ABSSWEB.spool/US10664025/runat_08052006_173459_28431/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02p -USER=US10664025@CGN_1_1_101 @runat_08052006_173459_28431
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/PTUS COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	799	96.7	153	2	US-09-621-976-3903
2	789	95.5	158	2	US-09-621-976-5222
3	785	95.0	158	2	US-10-104-047-3674
4	429	51.9	106	2	US-09-621-976-5223
5	223.5	27.1	216	2	US-09-949-016-11232
6	220	26.6	56	2	US-09-621-976-6447
7	77	9.3	358	1	US-08-558-823-19
8	77	9.3	358	2	US-09-723-806A-8
9	76	9.2	158	2	US-09-583-110-3248
10	76	9.2	165	2	US-09-107-433-3954
11	76	9.2	323	2	US-09-489-039A-10162
12	75.5	9.1	385	2	US-09-328-352-7887

13	75.5	9.1	424	1	US-08-484-993B-18
14	75.5	9.1	424	1	US-08-484-158B-18
15	75.5	9.1	424	1	US-08-484-596A-18
16	75.5	9.1	424	1	US-08-480-150A-18
17	75.5	9.1	424	2	US-08-458-731-18
18	75.5	9.1	424	2	US-08-149-223A-18
19	75	9.1	1105	2	US-09-949-016-8227
20	75	9.1	1117	2	US-09-949-016-6148
21	74.5	9.0	424	1	US-08-453-472-8
22	74.5	9.0	424	1	US-08-038-948-2
23	74.5	9.0	424	1	US-08-038-948-5
24	74.5	9.0	424	1	US-08-453-952-8
25	74.5	9.0	424	1	US-08-862-903-8
26	74	9.0	290	2	US-09-655-908-6
27	74	9.0	290	2	US-09-655-908-8
28	72.5	8.8	156	2	US-09-270-767-36988
29	72.5	8.8	156	2	US-09-270-767-52205
30	72.5	8.8	432	2	US-09-393-245-2
31	71.5	8.7	649	2	US-10-104-047-3492
32	71.5	8.7	1020	1	US-08-070-301-3
33	71	8.6	245	2	US-09-461-325-367
34	71	8.6	245	2	US-10-012-542-367
35	71	8.6	245	2	US-10-115-123-367
36	71	8.6	288	2	US-09-949-016-10862
37	71	8.6	424	1	US-08-453-472-7
38	71	8.6	424	1	US-08-038-948-4
39	71	8.6	424	1	US-08-453-952-7
40	71	8.6	424	1	US-08-862-903-7
41	71	8.6	424	1	US-08-484-158B-61
42	71	8.6	860	2	US-10-290-579A-252
43	70.5	8.5	519	2	US-09-312-183A-3
44	70.5	8.5	528	2	US-09-489-039A-13077
45	70.5	8.5	536	2	US-09-449-632-2

ALIGNMENTS

RESULT 1

US-09-621-976-3903
; Sequence 3903, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3903
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-621-976-3903

Alignment Scores:

Pred. No.: 2.44e-100 Length: 153
Score: 799.00 Matches: 153
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.7% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-09-621-976-3903 (1-153)

Qy 1 ATGGCGTTAGAGTCTTCGCTGCTCTCGCTGCTTGTGATTGGACCGGTGCTCAGAACCTC 60

Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIletrpThrGlyAlaGluAsnLeu 20

61	CATGTGAAAATTAAGTTGCTCTCTCGAGCTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA	120
Qy		
21	HieVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
Db		
121	ACAGAAATCTCTATATATTTCCGAGTAAATTACATCTGGGAATGGCTGCCCTGCAAT	180
Qy		
41	SerArgAsnLeuTyrllePheAlaAspGluLeuHisIleuGlyMetGlyCysProAlaAsn	60
Db		
181	CGGATACATACATATGTATGAGTTTATATATCTTGTCTGATTTGGCATCAGGACA	240
Qy		
61	ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr	80
Db		
241	AGGTAGTTTTCAGAGAAACTCTCTCTTTTCAACCGAGCTGACTTTACCCCAAGAAAT	300
Qy		
81	ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn	100
Db		
301	ATAGATCATGACCCCTCAGGAATCCATTGAGGTGTTCCACCTCTAGGAATCAGGTGG	360
Qy		
101	IleAspHisAspProGlnGluIleHisIleuGluCysSerThrSerArgIysSerValTrp	120
Db		
361	CTTACACCAAGTTTCTACTCAGAAATCAAAATAAAATTTGGATCCTAGTCCTTTTATTGCTGAC	420
Qy		
121	LeuThrProValSerThrGluAsnGluIleIysLeuAspProSerProPheIleAlaAsp	140
Db		
421	TTTCAGACAACGACGAGAGTTAGGATTTAATCTTC	459
Qy		
141	PheGlnThrThrAlaGluGluLeuGlyLeuLeuIlePhe	153
Db		

RESULT 2

```

US-09-621-976-5222
; Sequence 5222, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Joibert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5222
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-621-976-5222

```

Alignment Scores:		
Pred. No.:	5,81e-99	Length:
Score:	789.00	Matches:
Percent Similarity:	100.0%	Conservative:
Best Local Similarity:	100.0%	Mismatches:
Query Match:	95.5%	Indels:
DB:	2	Gaps:
		0
		0
		158

US-10-664-025-43 COPY 53 511 (1-459) x US-09-621-976-5222 (1-158)

Qy	1	ATGGCGTTAGAAAGTCTTGATGCTCTCTCGCTGTCCTTGATTTGGACCGGTGCTGAGAACCTC	60
Db	1	MetAlaLeuGluValIleuMetLeuLeuAlaValLeuIletrpThrGlyAlaGluAsnLeu	20
Qy	61	CATGTGAAAATAAGTTGGTCTCTCGACATGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA	120
Db	21	HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
Qy	121	AGCAGAAATCTGTATATATTTTCGGGATCAAATTTACATCTGGGAATGGGCTCCCTGCAAAAT	180
Db	41	SerAspAspLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGluCysProAlaAsn	60

Qy	181	CGGATACATCATATGATATATGAGTTTATATATCTGTTCGTGATTTGGCATCAGGACA	240
Db	61	ArgIleHisThrTyrValTyGluPheIleTyrLeuValHrGAspCysGlyIleArgThr	80
Qy	241	AGGGTAGTTTCTGAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT	300
Db	81	ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn	100
Qy	301	ATAGATCATGACCTCAGGAAATCATTTGGAGTGTTCACCTCTTAGGAAATCAGTGTGG	360
Db	101	IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIlysserValTtp	120
Qy	361	CTTATACACAGTTTCTACTGAGAAATCAATAAAATTTGGATCCTAGTCTCTTTTATTGCTGAC	420
Db	121	LeuThrProValSerThrGluAsnGluIleIlyssLeuAspProSerProPheIleAlaAsp	140
Qy	421	TTTCAGACAACAGCAGAGAGCTAGGATTATTA	453
Db	141	PheGlnThrThrAlaGluGluLeuGlyLeuLeu	151

3. T. J. P. 3

```

RES001.3
US-10-104-047-3674
; Sequence 3674, Application US/10104047
; Patent NO. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4996
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674

```

Alignment Scores:		
Pred. No.:	2,05e-98	Length:
Score:	785.00	Matches:
Percent Similarity:	99.3%	Conservative:
Best Local Similarity:	99.3%	Mismatches:
Query Match:	95.0%	Indels:
DB:	2	Gaps:
		158
		150
		0
		1
		0
		0

US-10-664-025-43 COPY 53 511 (1-459) x US-10-104-047-3674 (1-158)

Qy	1	ATGCGTTAGAA	GCTTGA	TGCTCT	CGCTGCT	TGTGAT	TGGAC	CGGTGCT	GAGAAC	CTTC	60	
Db	1	MetAla	LeuGlu	ValMet	LeuLeu	AlaVal	LeuIle	TrpThr	GlyAla	GluAsn	Leu 20	
Qy	61	CATCTGAA	ATAA	TAGT	CTCT	CTCG	ACTGGT	TGTAG	TGGT	CTCAG	TATCCAGT	TGCAGAA 120
Db	21	HisVal	IleSer	CysSer	LeuAsp	TrpLeu	MetAla	SerVal	IlePro	ValAla	Glu 40	
Qy	121	AGCAGA	AAATCT	GTATAT	TATTTG	CGGATG	AAATAC	ATCT	TGGGA	TGGCGT	CCCTG	CCAAAT 180
Db	41	SerArg	AsnLeu	TyrIle	PheAla	AspGlu	LeuHis	LeuGly	MetGly	CysPro	AlaAsn 60	
Qy	181	CGGATAC	ATACAT	ATATAT	TATG	AGTTAT	TATAT	CTTGT	TTCGT	GATTTG	GGCATC	CAGGACA 240
Db	61	ArgIle	HisThr	TyrVal	TyrGlu	PheIle	TyrLeu	ValArg	AspCys	GlyIle	ArgThr 80	
Qy	241	AGGTAGT	TTCTG	AGGAA	CTCT	CTCTTT	TTCA	ACCG	AGCTGT	CTACTT	TACCCCA	AGGAAT 300
Db	81	ArgVal	ValSer	GluGlu	ThrLeu	LeuPhe	GlnThr	GluLeu	TyrPhe	ThrPro	ArgAsn 100	
Qy	301	ATAGAT	CATGAC	CCCTC	AGGAA	ATCC	ATTG	GAGTGT	TCCAC	CTCT	TAGGAA	ATCAGTGTGG 360

Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120

QY 361 CTTACACCACTTCTACTGAGATGAATAAATTCGATCTTATGCTGCTGAC 420

Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140

QY 421 TTTCCAGACACAGCAGAGAGTTAGGATTATTA 453

Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 4

US-09-621-976-5223

; Sequence 5223, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 5223

; LENGTH: 106

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: SIGNAL

; LOCATION: -17..-1

US-09-621-976-5223

Alignment Scores:

Pred. No.: 9.5e-50 Length: 106

Score: 429.00 Matches: 82

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 51.9% Indels: 0

DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-09-621-976-5223 (1-106)

QY 1 ATGCGCTTGAAGTCTTGATGCTCTCGTCTTATGACCGGCTGAGACCTC 60

Db 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20

QY 61 CATGTGAAATAAGTTGCTCTCGGACTGGTCTCATGCTCTCAGTTATCCAGTTGCAGAA 120

Db 21 HisValValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40

QY 121 AGCAGAAATCTGATATATTTGCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 180

Db 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60

QY 181 CGGATACATACATATGATATGATTTATATATCTTGTTCGCGATGTCGATCAGGACA 240

Db 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80

QY 241 AGGGTA 246

Db 81 ArgVal 82

RESULT 5

US-09-949-016-11232

; Sequence 11232, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11232

; LENGTH: 216

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-11232

Alignment Scores:

Pred. No.: 1.85e-21 Length: 216

Score: 223.50 Matches: 40

Percent Similarity: 60.0% Conservative: 32

Best Local Similarity: 33.3% Mismatches: 43

Query Match: 27.1% Indels: 5

DB: 2 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x US-09-949-016-11232 (1-216)

QY 13 GTCTTGATGCTCTCTGCTGCTCTTGTGACCGGTGCTGAGAACCTCCATGTGAAATA 72

Db 13 LeuMetIleLeuLeuThrSerAlaPheSerAlaGlySerGlyGlnSerProMetThrVal 32

QY 73 AGTTGCTCTCTGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGAAATCTG 132

Db 33 LeuCysSerIleAspTrpPheMetValThrValHisProPheMetLeuAsnAspVal 52

QY 133 TATATATTTGCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAATCGGATACATACA 192

Db 53 CysValHisPheHisGluLeuHisLeuGlyLeuGlyCysProHisHisValGlnPro 72

QY 193 TATGATATGAGTTTATATATCTTGTCTGATTTGTCGATGTCAGGACAGGGTAGTTTCT 252

Db 73 HisAlaTyrlleGlnPheThrTyrlleArgValThrGluCysGlyIleArgAlaLysAlaValSer 92

QY 253 GAGGAAACTCTCTTTTCAACCCGAGCTGTACTTTTACCCCAAGGAATATAGATCATGAC 312

Db 93 GlnAspMetValIleTyrlleSerThrGluIleHisTyrlleSerSerLysGly-----Thr 109

QY 313 CCTCAGGAA-----ATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGGCTTACA 366

Db 110 ProSerLysPheValIleProValSerCysAlaAlaProGlnLysSerProTrpLeuThr 129

RESULT 6

US-09-621-976-6447

; Sequence 6447, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 6447

; LENGTH: 56

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-621-976-6447

Alignment Scores:

Pred. No.: 2.91e-21 Length: 56

Score: 220.00 Matches: 48

Percent Similarity: 98.0% Conservative: 0

```
Best Local Similarity: 98.0% Mismatches: 1
Query Match: 26.6% Indels: 1
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-09-621-976-6447 (1-56)

QY 308 ATGACCTCAGAAATCCATTGGAG-TGTTCCACCTCTAGGAAATCAGTGTGGCTTACA 366
Db 1 MetThrLeuArgLysSerIleTrpGluCysSerThrSerArgLysSerValTrpLeuThr 20
QY 367 CCAGTTCTTACTGAGAAATGAATAATGGATCCTAGTCCCTTTATTGCTGCACTTCAG 426
Db 21 ProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAspPheGln 40
QY 427 ACAACAGCAGAAAGATTAGATTATTA 453
Db 41 ThrThrAlaGluGluLeuGlyLeuLeu 49

RESULT 7
US-08-558-823-19
; Sequence 19, Application US/08558823
; Patent No. 5876994
; GENERAL INFORMATION:
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Roelofs, Wendell L.
; APPLICANT: Miller, Stuart J.
; TITLE OF INVENTION: PHEROMONE DESATURASES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,823
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rogalekyj Esq., Peter
; REGISTRATION NUMBER: 38,601
; REFERENCE/DOCKET NUMBER: 19603/400
; TELEPHONE: (716) 263-1634
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-558-823-19

Alignment Scores:
Pred. No.: 0.272 Length: 358
Score: 77.00 Matches: 38
Percent Similarity: 39.5% Conservativeness: 22
Best Local Similarity: 25.0% Mismatches: 50
Query Match: 9.3% Indels: 42
DB: 1 Gaps: 8

US-10-664-025-43_COPY_53_511 (1-459) x US-08-558-823-19 (1-358)

QY 124 AGAAATCTGTATATATTTCCGATCAATTTACATCTGGGA-----ATGGGC 168
Db 73 ArgAsnIleIleLeuMetAla---LeuLeuHisValGlyAlaLeuTyrGlyIleThrLeu 91
QY 169 TGCCCTGCAAAATCGGATACATATATATATGAG-----TTTATATATCTTGTTCGTGAT 225
Db 92 IleProSerSerLysValTyrThrLeuLeuTrpGlyIlePheTyrTyrLeuIleSerAla 111
QY 226 TGTGCATC----- 234
Db 112 LeuGlyIleThrAlaGlyAlaHisArgLeuTrpSerHisArgThrTyrLysAlaArgLeu 131

US-09-723-806A-8
; Sequence 8, Application US/09723806A
; Patent No. 6686185
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Logan, Thomas Joseph
; TITLE OF INVENTION: 25934, A NOVEL FATTY ACID DESATURASE AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 10448-024001
; CURRENT APPLICATION NUMBER: US/09/723,806A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-723-806A-8

Alignment Scores:
Pred. No.: 0.272 Length: 358
Score: 77.00 Matches: 38
Percent Similarity: 39.5% Conservativeness: 22
Best Local Similarity: 25.0% Mismatches: 50
Query Match: 9.3% Indels: 42
DB: 2 Gaps: 8

US-10-664-025-43_COPY_53_511 (1-459) x US-09-723-806A-8 (1-358)

QY 124 AGAAATCTGTATATATTTCCGATCAATTTACATCTGGGA-----ATGGGC 168
Db 73 ArgAsnIleIleLeuMetAla---LeuLeuHisValGlyAlaLeuTyrGlyIleThrLeu 91
QY 169 TGCCCTGCAAAATCGGATACATATATATATGAG-----TTTATATATCTTGTTCGTGAT 225
Db 92 IleProSerSerLysValTyrThrLeuLeuTrpGlyIlePheTyrTyrLeuIleSerAla 111
QY 226 TGTGCATC----- 234
Db 112 LeuGlyIleThrAlaGlyAlaHisArgLeuTrpSerHisArgThrTyrLysAlaArgLeu 131
```


Qy 235 -----AGGACAAGGTTAGTTTCTGTAGGAAACTCTCCCTTTTCAAACCGAGTGTACTTT 288
 Db 132 ProLeuArgIlePheLeuIleAlaAsnThrMetAlaPheGlnAspValTyrGlu 151
 Qy 289 ACCCAAGGAAT-----ATAGATCATGACCTCAGGAAATCCATTTCGAGTGTGTTCCACC 342
 Db 152 TrpAlaArgAspHisArgAlaHisHisLysPheSerGluThrHisAlaAspProHisAsn 171
 Qy 343 TCTAGGAAATCAGTG-----TGGCTTACA-----CCAGTT 372
 Db 172 SerArgArgGlyPhePhePheSerHisValGlyTrpLeuLeuValArgLysHisProAla 191
 Qy 373 TCTACTCAGATGAATAAATAATTCGATCCTAGTCTTTATTCGTGAC----- 420
 Db 192 ValLysGluLysGlyLysLeuAspMetSerAspLeuLysAlaGluLysLeuValMet 211
 Qy 421 TTTCCAGACAACGACAGAGAGCTAGGATTATTAAATC 456
 Db 212 PheGlnArgGlyTyrLysProGlyLeuLeuLeu 223

RESULT 9

US-09-583-110-3248
 ; Sequence 3248, Application US/09583110
 ; Patent No. 6699703
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al.
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 ; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
 ; FILE REFERENCE: PATH00-07A
 ; CURRENT APPLICATION NUMBER: US/09/583,110
 ; CURRENT FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/107,433
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/085,131
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: US 60/051,553
 ; PRIOR FILING DATE: 1997-07-02
 ; NUMBER OF SEQ ID NOS: 5322
 ; SEQ ID NO 3248
 ; LENGTH: 158
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 US-09-583-110-3248

Qy 304 GATCATGACCTCAGGAATC-----CAATTGGAGTGTTCACCTTAGGAAATCAGTG 357

Db ::: :::||||| ||||| ||| ::: :::

92 GluGluGluGluGluGluGluIleGluArgLeuGluGlnSerLysAlaAsnGlnLeuPhe 111

Qy 358 TGG-----CCTTACCACGTTTCTACTGAGAATGAATAAATTGGATCCTAGT 405

Db ||| ||||| ||| ::: ::: :::

112 TrpThrValThrLeuLeuThrIleAlaSerGlyGlyAspAsnLeuGlyIle---TyrIle 130

Qy 406 CTTTTATTGTGCTACTTTCACACACACACAGAGTAGGTAGGATTTAATCTTC 459

Db ||| ||| ::: ::: ||||| ::: |||

131 ProTyrPheAlaSerLeuAspTrpSerGlnThrLeuValAlaLeuLeuValPhe 148

RESULT 10

US-09-107-433-3954

; Sequence 3954, Application US/09107433

; Patent No. 6800744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

 MEDIUM TYPE: CD-ROM ISO9660

 COMPUTER: <Unknown>

 OPERATING SYSTEM: <Unknown>

 SOFTWARE: <Unknown>

 CURRENT APPLICATION DATA:

 APPLICATION NUMBER: US/09/107,433

 FILING DATE: 30-Jun-1998

 PRIOR APPLICATION DATA:

 APPLICATION NUMBER: 60/ 085131

 FILING DATE: May 12, 1998

 APPLICATION NUMBER: 60/051553

 FILING DATE: July 2, 1997

 ATTORNEY/AGENT INFORMATION:

 NAME: Ariniello, Pamela Deneke

 REGISTRATION NUMBER: 40,489

 REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

 TELEPHONE: (781)893-5007

 TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3954:

SEQUENCE CHARACTERISTICS:

 LENGTH: 165 amino acids

 TYPE: amino acid

 TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

 NAME/KEY: misc feature

 LOCATION: (B) LOCATION 1...165

SEQUENCE DESCRIPTION: SEQ ID NO: 3954:

US-09-107-433-3954

Alignment Scores:		
Pred. No.:	0.256	165
Score:	76.00	39
Percent Similarity:	41.8%	27
Best Local Similarity:	24.7%	Conservative: 46
Query Match:	9.2%	Mismatches: 46
DB:	2	Indels: 9
		Gaps: 9

```

Alignment Scores:
Pred. No.:      0.251      Length:      158
Score:          76.00      Matches:      39
Percent Similarity: 41.8%      Conservaive: 27
Best Local Similarity: 24.7%      Mismatches: 46
Query Match:      9.2%      Indels:      46
DB:              2      Gaps:      9

US-10-664-025-43_COPY_53_511 (1-459) x US-09-583-110-3248 (1-158)

QY      64  GTGAAATAAGTTGCTCTCTGACGCTGGTGTGATCTCAGTTATCCAGTTCCAGAA--- 120
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      11  ValTyrIleSerThrSerIleAspTyrLeuIleIleLeuIleLeuPheAlaGlnLeu 30
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      121 -----ACGAGAAATCTGTATATATTTCCGGATGAATTCATCTGGGAATGGGC--- 168
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| |||
Db      31  SerGlnAsnLysGlnTyrPheIleTyrAlaGlyGln---TyrLeuGlyThrGlyLeu 49
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| |||
QY      169 ---TGCCTCGAAATCCGATACATACATATCTATATCAGTTT----- 207
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| |||
Db      50  LeuValGlyAlaSerLeuValAlaAlaTyrValValAsnPheValProGluGluTyrMet 69
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| |||
QY      208 -----ATATATCTTCTCGTATTGTGGCATCCAGCAAGG 243
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      70  ValGlyLeuLeuGlyLeuIleProIleTyrLeu-----GlyIleArgPheAla 85
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      244  GTAGTTTCTGAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAATATA 303
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      86  IleValGlyGluAspAla-----GlyIleArgPheAla 85
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db      190 SerAlaThrGlu 193
RESULT 13
US-08-484-993B-18
; Sequence 18, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-993B-18
Alignment Scores:
Pred. No.: 0.473
Score: 75.50
Percent Similarity: 37.1%
Best Local Similarity: 22.6%
Query Match: 9.1%
DB: 1
US-10-664-025-43_COPY_53_511 (1-459) x US-08-484-993B-18 (1-424)
QY      1 ATGGCGTTAGAACTTGTGATCTCGCTGCTTGTGATTTGACCGGTGCT-----51
Db      1 MetGlyLeuSerTyrGlyLeuPheIleCysPheLeuLeuTrpAlaGlyThrGlyLeuCys 20
QY      52 -----GAGAACCTCCAT-----GTG 66
Db      21 TyrProThrThrThrGluAspLysThrHisProSerLeuProSerSerProSerVal 40
QY      67 AAAATAAGTTGCTCTCTGACCTGGTTGATGGTCTCAGTTATCCAGTTCCAGAAAGCAGA 126
Db      41 ValValGluCysArgHisAlaTrpLeuValValAsnVal-----SerLys 55
QY      127 ATCTGTAT-----ATATTTGGGATGAATTACATCTGGGAATGGCTGC 171
Db      56 AsnLeuPheGlyThrGlyArgLeuValArgProAlaAspLeuThrLeuGly-----72
QY      172 CTGCAAAAT-----CGGATACATACA 192
Db      73 ProGluAsnCysGluProLeuIleSerGlyAspSerAspThrValArgPheGluVal 92
QY      193 TATGTATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACAAGGGTAGTTCT 252
Db      93 GluLeuHisLys-----CysGlyAsnSerValGln-----ValThr 104
QY      253 GAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAATATAGATCAT---309
Db      105 GluAspAlaLeuValTyrSerThrPheLeuLeuHisAsnProArgProMetGlyAsnLeu 124
QY      310 -----GACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCA 354
Db      125 SerIleLeuArgThrAsnArgAlaGluValProIleGluCysArgTyrProArgHisSer 144
QY      355 GTGTGGCTTACACCGATTTCTACTGAGATGAATAAATTGGATCCTAGTCTTTTATT 414
Db      145 -----AsnValSerSerGluAlaIleLeu-----ProThrTrp 155
QY      415 GCTGACTTTCAGACAACA 432
Db      156 ValProPheArgThrThr 161
RESULT 14
US-08-484-158B-18
; Sequence 18, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immunocontraception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:

```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-158B-18

Alignment Scores:
Pred. No.: 0.473 Length: 424
Score: 75.50 Matches: 42
Percent Similarity: 37.1% Conservative: 27
Best Local Similarity: 22.6% Mismatches: 50
Query Match: 9.1% Indels: 67
DB: 1 Gaps: 11

US-10-664-025-43_COPY_53_511 (1-459) x US-08-484-158B-18 (1-424)
QY 1 ATGGCGTTAGAGTCTTGATGCTCGCTGCTGTTGATTGGACCGGTGCT----- 51
Db 1 MetGlyLeuSerTyrglyLeuPheileCysPheLeuLeuTrpAlaGlyThrGlyLeuCys 20
QY 52 -----GAGAACCTCCAT-----GTG 66
Db 21 TyrProProThrThrGluAspLysThrHisProSerLeuProSerSerProSerVal 40
QY 67 AAATAAGTTGCTCTCTGAGCTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGA 126
Db 41 ValValGluCysArgHisAlaTrpLeuValValAsnVal-----SerLys 55
QY 127 AATCTGTAT-----ATATTTGCGGATGAATTACATCTGGGAATGGGCTGC 171
Db 56 AsnLeuPheGlyThrGlyArgLeuValArgProAlaAspLeuThrLeuGly----- 72
QY 172 CTGCAAAAT-----CGGATACATACA 192
Db 73 ProGluAsnCysGluProLeuileSerGlyAspSerAspThrValArgPheGluVal 92
QY 193 TATGTATATAGTTTTATATATCTTCTGTCGATTCGTCATCAGGACAGGAGGTAGTTCT 252
Db 93 GluLeuHisLys-----CysGlyAsnSerValGln---ValThr 104
QY 253 GAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAATATAGATCAT--- 309
Db 105 GluAspAlaLeuValTySerThrPheLeuLeuHisAsnProArgProMetGlyAsnLeu 124
QY 310 -----GACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCA 354
Db 125 SerIleLeuArgThrAsnArgAlaGluValProIleGluCysArgTyrProArgHisSer 144
QY 355 GTGTGCTTACACCGATTTCTACTCAGATGAATAAAATTCGATCCTAGTCTTTTATT 414
Db 145 -----AsnValSerSerGluAlaIleLeu-----ProThrTrp 155
QY 415 GCTGACTTTACAGAACACA 432
Db 156 ValProPheArgThrThr 161

RESULT 15
US-08-484-596A-18
; Sequence 18, Application US/08484596A
; Patent No. 5981228
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Heu, Kuang T.
; APPLICANT: Podolaki, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
```

QY 355 GTGTGGCTTACACCGTTTCTACTGAGATGAATAAAATTGGATCCTAGTCCTTTTATT 414
 Db 145 -----AsnValSerSerGluAlaIleLeu-----ProThrTrp 155
 QY 415 GCTGACTTTCAGACAACA 432
 Db 156 ValProPheArgThrThr 161

Search completed: May 9, 2006, 10:10:08
 Job time : 37.1 secs

This Page Blank (Inten)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:09:27 ; Search time 16.29 Seconds
(without alignments)
3531.927 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511

Perfect score: 826

Sequence: 1 atggcgttagaagctctgat.....agttaggattatacttc 459

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10664025/runat_08052006_173502_28483/app_query.fasta_1
-DB=Published Applications AA_Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss07
-USER=US10664025 @CGN 1 1_198 @runat_08052006_173502_28483 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA_Main.*

1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	789	95.5	158	3	US-09-731-872-260
2	789	95.5	158	3	US-09-876-997-260
3	789	95.5	158	4	US-10-050-704-113
4	789	95.5	158	4	US-10-467-535-4
5	789	95.5	158	4	US-10-798-512-113
6	789	95.5	158	5	US-10-467-046-3
7	789	95.5	158	5	US-10-843-836-260
8	789	95.5	158	5	US-10-472-533-464
9	785	95.0	158	4	US-10-104-047-3674
10	429	51.9	106	3	US-09-731-872-265
11	429	51.9	106	3	US-09-876-997-265

12	429	51.9	106	5	US-10-643-836-265	Sequence 265, App
13	225.5	27.3	212	4	US-10-287-971-40	Sequence 40, Appl
14	223.5	27.1	212	3	US-09-745-763-203	Sequence 203, App
15	223.5	27.1	212	4	US-10-047-021-91	Sequence 91, Appl
16	223.5	27.1	212	5	US-10-970-493-91	Sequence 91, Appl
17	223.5	27.1	213	4	US-10-411-224-91	Sequence 91, Appl
18	161	19.5	163	4	US-10-287-971-42	Sequence 42, Appl
19	98.5	11.9	930	5	US-10-450-763-42407	Sequence 42407, A
20	94	11.4	59	3	US-09-864-761-36906	Sequence 36906, A
21	88.5	10.7	234	5	US-10-450-763-57701	Sequence 57701, A
22	85.5	10.4	5373	4	US-10-408-765A-741	Sequence 741, App
23	85.5	10.4	5430	5	US-10-805-684-151	Sequence 151, App
24	84	10.2	1241	4	US-10-282-122A-46016	Sequence 46016, A
25	78.5	9.5	500	5	US-10-527-556-382	Sequence 382, App
26	77	9.3	358	4	US-10-377-072-14	Sequence 14, Appl
27	77	9.3	358	4	US-10-377-072-14	Sequence 14, Appl
28	76	9.2	109	4	US-10-425-115-309950	Sequence 309950,
29	76	9.2	165	5	US-10-617-320-3954	Sequence 3954, Ap
30	76	9.2	204	5	US-10-472-928-3336	Sequence 3336, Ap
31	76	9.2	2957	5	US-10-732-923-8692	Sequence 8692, Ap
32	75.5	9.1	403	5	US-10-732-923-11647	Sequence 11647, A
33	75	9.1	289	4	US-10-282-122A-52488	Sequence 52488, A
34	75	9.1	1117	4	US-10-788-792-238	Sequence 238, App
35	75	9.1	1117	5	US-10-756-149-5461	Sequence 5461, Ap
36	75	9.1	1118	3	US-09-909-567B-42	Sequence 42, Appl
37	75	9.1	1120	5	US-10-450-763-40643	Sequence 40643, A
38	74.5	9.0	384	4	US-10-369-493-2111	Sequence 2111, Ap
39	74.5	9.0	525	4	US-10-425-115-253855	Sequence 253855,
40	74	9.0	290	4	US-10-700-171-6	Sequence 6, Appli
41	74	9.0	290	4	US-10-437-963-147028	Sequence 147028,
42	73.5	8.9	832	4	US-10-087-192-9	Sequence 9, Appli
43	73	8.8	376	4	US-10-767-701-35932	Sequence 35932, A
44	72.5	8.8	233	4	US-10-205-331-42	Sequence 42, Appl
45	72	8.7	358	4		

ALIGNMENTS

RESULT 1

US-09-731-872-260
; Sequence 260, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Mine Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-731-872-260

Alignment Scores:
Pred. No.: 1.02e-88
Score: 789.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 95.5%
DB: 3
Length: 158
Matches: 151
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-09-731-872-260 (1-158)

```
QY 1 ATGGCGTTAGAGTCTTGTATGCTCCTCGCTGCTTGTGACCGGTGCTCAGAACCTC 60
    |||||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
    |||||||
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
    |||||||
Db 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
    |||||||
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAA 180
    |||||||
Db 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
    |||||||
QY 181 CGGATACATACATGATATATAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 240
    |||||||
Db 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80
    |||||||
QY 241 AGGCTAGTTTCTGAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
    |||||||
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
    |||||||
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
    |||||||
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
    |||||||
QY 361 CTTACACAGTTTCTACTCAGAAATCAATAAAATTTGGATCCCTAGTCTTTTATTGCTGAC 420
    |||||||
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
    |||||||
QY 421 TTTTCAGACACAGCAGAAAGTTAGGATTATTA 453
    |||||||
Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
    |||||||
```

RESULT 2

```
US-09-876-997-260
; Sequence 260, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Mine Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US4 CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-876-997-260
```

```
Alignment Scores:
Pred. No.: 1,02e-88 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 3 Gaps: 0
```

US-10-664-025-43_COPY_53_511 (1-459) x US-09-876-997-260 (1-158)

```
QY 1 ATGGCGTTAGAGTCTTGTATGCTCCTCGCTGCTTGTGACCGGTGCTCAGAACCTC 60
    |||||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
    |||||||
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
    |||||||
Db 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
    |||||||
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAA 180
    |||||||
Db 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
    |||||||
QY 181 CGGATACATACATGATATATAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 240
    |||||||
Db 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80
    |||||||
QY 241 AGGCTAGTTTCTGAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
    |||||||
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
    |||||||
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 360
    |||||||
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
    |||||||
QY 361 CTTACACAGTTTCTACTCAGAAATCAATAAAATTTGGATCCCTAGTCTTTTATTGCTGAC 420
    |||||||
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
    |||||||
QY 421 TTTTCAGACACAGCAGAAAGTTAGGATTATTA 453
    |||||||
Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
    |||||||
```

RESULT 3

```
US-10-050-704-113
; Sequence 113, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039FI
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-113
```

```
Alignment Scores:
Pred. No.: 1,02e-88 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 4 Gaps: 0
```

US-10-664-025-43_COPY_53_511 (1-459) x US-10-050-704-113 (1-158)

```
QY 1 ATGGCGTTAGAGTCTTGTATGCTCCTCGCTGCTTGTGACCGGTGCTCAGAACCTC 60
    |||||||
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
    |||||||
QY 61 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 120
    |||||||
```



```

Db      21 HisVallysileSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Qy      121 AGCAGAAATCTGTATATATTTCCGAGTGAATTAATCATCTGGGAATGGGCTGCCCTGCAAAAT 180
Db      41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Qy      181 CGGATACATACATATGTATATAGTTTATATATCTTGTCTGCTGATTCTGGATTCGGCATCAGGACA 240
Db      61 ArgIleHisThrTyValTyGluPheIleTyrlleTyrlleValArgAspCysGlyIleArgThr 80
Qy      241 AGGTAGTCTTCTGAGGAAATCTCTCTTTTCAAAACCGAGCTGTACTTTTACCCCAAGGAAT 300
Db      81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
Qy      301 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAATCAGTGTGG 360
Db      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Qy      361 CTTACACAGTTTCTACTGAGAATGAATAAATTGGATCCTAGTCTCTTTTATTGCTGAC 420
Db      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Qy      421 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 453
Db      141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151

RESULT 4
US-10-467-535-4
; Sequence 4, Application US/10467535
; Publication No. US20040146970A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; YAO, Monique G.
; APPLICANT: ISON, Craig H.; LU, Yan
; APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
; APPLICANT: BAUGHN, Mariah R.; DING, Li
; APPLICANT: XU, Yuming; GIETZEN, Kimberly J.
; APPLICANT: TANG, Tom Y.; LAL, Preeti G. Neil
; APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
; APPLICANT: LU, Dying Aina M.; RICHARDSON, Thomas W.
; APPLICANT: TRAN, Uyen K.; KHARE, Reena
; APPLICANT: CHAWLA, Narinder K.
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PE-0503 USN
; CURRENT APPLICATION NUMBER: US/10/467,535
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/US02/03715
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/268,111
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/271,175
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,503
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/274,552
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 5284076CD1
US-10-467-535-4

Alignment Scores:
Pred. No.: 1,02e-88 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-10-467-535-4 (1-158)
Qy      1 ATGGCGTTAGAAATCTTGTATGCTCGCTCGCTGTCTTGTATTTGGACCGGTGCTGAGAACCTC 60
Db      1 MetAlaLeuGluValLeuMetLeuLeuAlaValleuIleTrpThrGlyAlaGluAsnLeu 20
Qy      61 CATGTGAAATAAGTTGCTCTCTGGACTGTTGATGTCTCAGTATATCCAGTTTGCAGAA 120
Db      21 HisVallysileSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Qy      121 AGCAGAAATCTGTATATATTTCCGAGTGAATTAATCATCTGGGAATGGGCTGCCCTGCAAAAT 180
Db      41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Qy      181 CGGATACATACATATGTATATAGTTTATATATCTTGTCTGCTGATTCTGGATTCGGCATCAGGACA 240
Db      61 ArgIleHisThrTyValTyGluPheIleTyrlleTyrlleValArgAspCysGlyIleArgThr 80
Qy      241 AGGTAGTCTTCTGAGGAAATCTCTCTTTTCAAAACCGAGCTGTACTTTTACCCCAAGGAAT 300
Db      81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
Qy      301 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTCCACCTCTAGGAATCAGTGTGG 360
Db      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Qy      361 CTTACACAGTTTCTACTGAGAATGAATAAATTGGATCCTAGTCTCTTTTATTGCTGAC 420
Db      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Qy      421 TTTCAGACAACAGCAGAGAGTTAGGATTATTA 453
Db      141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151

RESULT 5
US-10-798-512-113
; Sequence 113, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 113
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-798-512-113

Alignment Scores:
Pred. No.: 1,02e-88 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-10-798-512-113 (1-158)
Qy      1 ATGGCGTTAGAAATCTTGTATGCTCGCTCGCTGTCTTGTATTTGGACCGGTGCTGAGAACCTC 60
```

Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAAAAAGTTTCTCTGACTGTTGATGCTCGCTCAGTTATCCAGTTGCAGAA 120
Db 21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTCGGAATGGCTGCCCTGCAAT 180
Db 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATATATATATATCTTGTTCGTGATGTGGCATCAGGACA 240
Db 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGGTAGTTTCTGAGGAACCTCTCTTTTCAAACCGAGCTGTACTTACCCCAAGGAAT 300
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIlySerValTyr 120
QY 361 CTTACACAGTTTCTACTGAGAATGAAATAAAATTGGATCCTAGTCTCTTTTATTCGTGAC 420
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTCAGACACAGCAGAGTAGGATTATTA 453
Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 6
US-10-467-046-3
; Sequence 3, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(17)
US-10-467-046-3

Alignment Scores:
Pred. No.: 1.02e-88 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-10-467-046-3 (1-158)
QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTCAGAACCTC 60
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20

QY 61 CATGTGAAAAAAGTTTCTCTGACTGTTGATGCTCGCTCAGTTATCCAGTTGCAGAA 120
Db 21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTCGGAATGGCTGCCCTGCAAT 180
Db 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATATATATATATCTTGTTCGTGATGTGGCATCAGGACA 240
Db 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGGTAGTTTCTGAGGAACCTCTCTTTTCAAACCGAGCTGTACTTACCCCAAGGAAT 300
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIlySerValTyr 120
QY 361 CTTACACAGTTTCTACTGAGAATGAAATAAAATTGGATCCTAGTCTCTTTTATTCGTGAC 420
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTCAGACACAGCAGAGTAGGATTATTA 453
Db 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 7
US-10-643-836-260
; Sequence 260, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-10-643-836-260

Alignment Scores:
Pred. No.: 1.02e-88 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-10-643-836-260 (1-158)

QY 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTCAGAACCTC 60
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAAAAAGTTTCTCTGACTGTTGATGCTCGCTCAGTTATCCAGTTGCAGAA 120

Db	21	HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
Qy	121	AGCAGAAATCTGTATATATTTCCGGATGAATTTACATCTCGGAATGGCTGCCTGCAAAT	180
Db	41	SerArgAsnLeuTyrllePheIlealaspGluLeuHisLeuGlyMetGlyCysProIleAsn	60
Qy	181	CGGATACATACATATGTATATGAGTTTATATATCTTGTTCGATTTGGCATCTGGCATCAGGACA	240
Db	61	ArgIleHisThrTyValTyrgluPheIleTyLeuValArgAspCysGlyIleArgThr	80
Qy	241	AGGGTAGTTCTCGAGAACTCTCTTTTCAACCGAGCTGTACTTTACCCAGGAAT	300
Db	81	ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn	100
Qy	301	ATAGATCATGACCCTCAGGAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG	360
Db	101	IleaspHisaspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp	120
Qy	361	CTTACACCAAGTTTCTACTCGAAGATGAAATTAATAATTCGGATCCCTVAGTCTTTTATTGCTGAC	420
Db	121	LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleIlealasp	140
Qy	421	TTTCAGACACAGCAGGAAGAGTTAGGATTATTA	453
Db	141	PheGlnThrThrAlaGluGluLeuGlyValLeuLeu	151

```

RESULT 8
US-10-472-533-464
; Sequence 464, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
;
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS906PCT
; CURRENT APPLICATION NUMBER: US/10/472,533
; CURRENT FILING DATE: 2003-09-20
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 650
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-472-533-464

```

Alignment Scores:	
Pred. No.:	1.02e-88
Score:	789.00
Length:	158
Matches:	151
Percent Similarity:	100.0%
Conservative:	0
Best Local Similarity:	100.0%
Mismatches:	0
Query Match:	95.5%
Indels:	0
DB:	5
Gaps:	0

US-10-664-025-43 COPY 53 511 (1-459) x US-10-472-533-464 (1-158)

1	ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGATTTGGACCGTGCTGGAACCTC	60
QY		
1	MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu	20
Db		
61	CATGTGAAATAAAGTTGCTCTCTGACATGGTTGATGGTCTCAGTTATCCAGTTGCAGAA	120
QY		
21	HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
Db		
121	ACGAGAAATCTGTATATATTTGCGATCAATTACATCTGGGAATGGCTCCCTCGCAAT	180
QY		
41	SerArgAsnLeuTyrlIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn	60
Db		
181	CGGATACATACATATGTATATGAGTTTATATATCTTGTTCTGTATCTGGCATCAGACA	240
QY		

[illegible]

```

RESULT 9
US-10-104-047-3674
; Sequence 3674, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674

```

Alignment Scores:			
Pred. No.:	3,228-88	Length:	158
Score:	785.00	Matches:	150
Percent Similarity:	99.3%	Conservative:	0
Best Local Similarity:	99.3%	Mismatches:	1
Query Match:	95.0%	Indels:	0
DB:	4	Gaps:	0

US-10-664-025-43 COPY 53 511 (1-459) x US-10-104-047-3674 (1-158)

1	ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCTGGAACCTC	60
1	MetaLiaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAenLeu	20
61	CATGTGAAAAATAAGTTGCTCTCTCGACTGGTTGATGCTCAAGTTATCCAGTTGCAGAA	120
21	HisValIlyIleSerCysSerLeuAspTrpLeuMetAlaSerValIleProValAlaGlu	40
121	AGCAGAAATCGTATATATTTGGGANGAATTACATCTGGGAATGGCTGCCCTGCAAA	180
41	SerArgAenLeuTyrlIlePheAlaaspGluLeuHisLeuGlyMetGlyCysProAlaAen	60
181	CGGATACATACATATGATATGAGTTTATATATCTCTTCGTCATTGTGGCATCAGACA	240
61	ArgIleHisThrTyValTyGluPheIleTyLeuValArgaspCysGlyIleArgThr	80
241	AGGTAAGTTTCTGAGGAAACTTCCTTTTTTCAACCGAGCTGTACTTTTACCCCAAGGA	300
81	ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAen	100
301	ATAGATCATGACCTCAGGAAATCCATTTGGAGTTTCCACCTCTTAGGAAATCAGTGTGG	360
101	IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTrp	120


```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-643-836-265

Alignment Scores:
Pred. No.: 5,44e-44 Length: 106
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 51.9% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-10-643-836-265 (1-106)

QY 1 ATGCGGTAGAGTCTTGTGCTCTCGCTGCTCTGATTTGGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuThrGlyAlaGluAsnLeu 20

QY 61 CATGTGAATAATAGTCTCTCTGACTGCTGATGCTGCTGATGCTGCTGATGCTGCA 120
DB 21 HisValIysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40

QY 121 AGCAGAACTGTATATATTGGGATGATTTACATCTGGGATGGGCTGCCCTGCNAAT 180
DB 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60

QY 181 CGGATACATACATATGATGATTTATATATCTTGTGCTGATTTGGCATCAGGACA 240
DB 61 ArgIleHisThrValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80

QY 241 AGGGTA 246
DB 81 ArgVal 82

RESULT 13
US-10-287-971-40
; Sequence 40, Application US/10287971
; Publication No. US20040067882A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-480A
; CURRENT APPLICATION NUMBER: US/10/287,971
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 09/997,425
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 10/035,568
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/338,626
; PRIOR FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/401,479
; PRIOR FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: 60/333,072
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/348,283
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/393,262
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/406,181
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 397
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 40
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-971-40

Alignment Scores:
Pred. No.: 1.42e-18 Length: 212
```

```

Score: 225.50 Matches: 40
Percent Similarity: 60.8% Conservative: 33
Best Local Similarity: 33.3% Mismatches: 42
Query Match: 27.3% Indels: 5
DB: 4 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x US-10-287-971-40 (1-212)

QY 13 GTCTTGATGCTCTCGCTGCTGCTGATTTGGACCGGTGCTGAGAACCTCCATGTGAAAATA 72
DB 9 LeuMetIleLeuLeuThrSerAlaLeuSerAlaGlySerGlyGlnSerProMetThrVal 28

QY 73 AGTTGCTCTCTGACTGGTTGAGTCTCAGTTATCCAGTTCCAGAAACGAGAAATCTG 132
DB 29 LeuCysSerIleAspTrpPheMetValThrValHisProPheMetLeuAsnAspVal 48

QY 133 TATATATTTGGCGATGAATTACATCTGGGAATGGGCTGCCCTGCAATCGGATACATACA 192
DB 49 CysValHisPheHisGluLeuHisLeuGlyLeuGlyCysProAsnHisValGlnPro 68

QY 193 TATGTATATAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACAAAGGAGTTTCT 252
DB 69 HisAlaTyrlleGlnPheThrTyrlleArgValThrGluCysGlyIleArgAlaLysAlaValSer 88

QY 253 GAGAAACTCTCTTTTCAACCGAGCTGTACTTTTACCCCAAGGATATAGATCATGAC 312
DB 89 GlnAspMetValIleTyrlleSerThrGluIleHisTyrlleSerSerLysGly-----Thr 105

QY 313 CCTCAGGAA-----ATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGCTTACA 366
DB 106 ProSerLysPheValIleProValSerCysAlaAlaProGlnLysSerProTrpLeuThr 125

RESULT 14
US-09-745-763-203
; Sequence 203, Application US/09745763
; Patent No. US20020065194A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; LaVallie, Edward R.
; Collins-Racie, Lisa A.
; Evans, Cheryl
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; ENCODING THEM
; NUMBER OF SEQUENCES: 219
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/745,763
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 203:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 203:
US-09-745-763-203

Alignment Scores:
Pred. No.: 2 51e-18 Length: 212
Score: 223.50 Matches: 40
Percent Similarity: 60.0% Conservative: 32
Best Local Similarity: 33.3% Mismatches: 43
Query Match: 27.1% Indels: 5
DB: 3 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x US-09-745-763-203 (1-212)
QY 13 GTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCTGAGAACCTCCATGTGAAATA 72
Db 9 LeuMetIleLeuLeuThrSerAlaPheSerAlaGlySerGlyGlnSerProMetThrVal 28
QY 73 AGTTGCTCTCTCGACTGGTTGATGCTCTCAGTTATCCCAAGTTCAGAGAGCAAAATCTG 132
Db 29 LeuCysSerIleAspTrpPheMetValThrValHisProPheMetLeuAsnAspVal 48
QY 133 TATATATTTGCGATGAAATTACATCTGGGAATGGCTGCCCTGCAAAATCGGATACATACA 192
Db 49 CysValHisPheHisGluLeuHisGlyLeuGlyCysProProAsnHisValGlnPro 68
QY 193 TATGTATATGAGTTTATATATCTTCTGTTGATGTCAGTTCAGGACCAAGGGTAGTTCT 252
Db 69 HisAlaTyrGlnPheThrArgValThrGluCysGlyIleArgAlaLysAlaValSer 88
QY 253 GAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAATATAGATCATGAC 312
Db 89 GlnAspMetValIleTyrSerThrGluIleHisTyrSerSerLysGly-----Thr 105
QY 313 CCTCAGGAA-----ATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGGCTTACA 366
Db 106 ProSerLysPheValIleProValSerCysAlaAlaProGlnLysSerProTrpLeuThr 125

RESULT 15
US-10-047-021-91
; Sequence 91, Application US/10047021
; Publication No. US20040002591A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: P2016P2
; CURRENT APPLICATION NUMBER: US/10/047,021
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,066
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: PCT/US98/18360
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,666
; PRIOR FILING DATE: 1997-09-12
```

```
; PRIOR APPLICATION NUMBER: US 60/090,112
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any amino acid
US-10-047-021-91

Alignment Scores:
Pred. No.: 2 51e-18 Length: 212
Score: 223.50 Matches: 40
Percent Similarity: 60.0% Conservative: 32
Best Local Similarity: 33.3% Mismatches: 43
Query Match: 27.1% Indels: 5
DB: 4 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x US-10-047-021-91 (1-212)
QY 13 GTCTTGATGCTCTCGCTGCTCTTGATTGGACCGGTGCTGAGAACCTCCATGTGAAATA 72
Db 9 LeuMetIleLeuLeuThrSerAlaPheSerAlaGlySerGlyGlnSerProMetThrVal 28
QY 73 AGTTGCTCTCTCGACTGGTTGATGCTCTCAGTTATCCCAAGTTCAGAGAGCAAAATCTG 132
Db 29 LeuCysSerIleAspTrpPheMetValThrValHisProPheMetLeuAsnAspVal 48
QY 133 TATATATTTGCGATGAAATTACATCTGGGAATGGCTGCCCTGCAAAATCGGATACATACA 192
Db 49 CysValHisPheHisGluLeuHisGlyLeuGlyCysProProAsnHisValGlnPro 68
QY 193 TATGTATATGAGTTTATATATCTTCTGTTGATGTCAGTTCAGGACCAAGGGTAGTTCT 252
Db 69 HisAlaTyrGlnPheThrArgValThrGluCysGlyIleArgAlaLysAlaValSer 88
QY 253 GAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAATATAGATCATGAC 312
Db 89 GlnAspMetValIleTyrSerThrGluIleHisTyrSerSerLysGly-----Thr 105
QY 313 CCTCAGGAA-----ATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGGCTTACA 366
Db 106 ProSerLysPheValIleProValSerCysAlaAlaProGlnLysSerProTrpLeuThr 125
```

Search completed: May 9, 2006, 10:13:14

Job time : 82.45 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:10:23 ; Search time 4.77 Seconds
(without alignments)
1336.144 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511

Perfect score: 826

Sequence: 1 atggcgctgaagctcttgat.....agttagattataattcttc 459

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 470810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DRV=xlp

-Q=/abss/ABSSWEB_spool/US10664025/runat_08052006_173505_28538/app_query.fasta.1

-DB=Published Applications AA.New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=2000000000 -HOST=abs02p

-USER=US10664025 @CGN 1 1 47 @runat_08052006_173505_28538 -NCPU=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA.New.*

1: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep1.*

2: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pep1.*

3: /SIDSS/ptodata/1/pubpaa/US07_NEW_PUB.pep1.*

4: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep1.*

5: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*

6: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*

7: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*

8: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep1.*

9: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep1.*

10: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep1.*

11: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep1.*

12: /SIDSS/ptodata/1/pubpaa/US60_NEW_PUB.pep1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	785	95.0	158	11	US-11-072-512-3674 Sequence 3674, Ap
2	775	93.8	148	9	US-10-475-075-788 Sequence 788, Ap
3	752	91.0	143	9	US-10-475-075-787 Sequence 787, Ap
4	223.5	27.1	216	9	US-10-821-234-1033 Sequence 1033, Ap
5	85.5	10.4	3803	9	US-10-995-561-773 Sequence 773, Ap

6	85.5	10.4	3960	9	US-10-995-561-771 Sequence 771, App
7	85.5	10.4	5335	9	US-10-995-561-777 Sequence 777, App
8	85.5	10.4	5406	9	US-10-995-561-774 Sequence 774, App
9	85.5	10.4	5415	9	US-10-995-561-779 Sequence 779, App
10	85.5	10.4	5464	9	US-10-995-561-775 Sequence 775, App
11	76.5	9.3	242	11	US-11-096-568A-8621 Sequence 8621, Ap
12	76.5	9.3	394	11	US-11-096-568A-8620 Sequence 8620, Ap
13	76.5	9.3	398	11	US-11-096-568A-8619 Sequence 8619, Ap
14	73	8.8	460	11	US-11-188-298-7840 Sequence 7840, Ap
15	71.5	8.7	440	11	US-11-045-004-81 Sequence 81, Appl
16	71.5	8.7	649	11	US-11-072-512-3492 Sequence 3492, Ap
17	70	8.5	954	9	US-10-453-372-1150 Sequence 1150, Ap
18	70	8.5	1268	9	US-10-453-372-1144 Sequence 1144, Ap
19	70	8.5	1268	9	US-10-453-372-1154 Sequence 1154, Ap
20	70	8.5	1288	9	US-10-453-372-1146 Sequence 1146, Ap
21	70	8.5	1288	9	US-10-453-372-1152 Sequence 1152, Ap
C 22	69.5	8.4	774	11	US-11-096-568A-33913 Sequence 33913, A
C 23	69.5	8.4	841	11	US-11-096-568A-33911 Sequence 33911, A
24	69	8.4	406	9	US-10-950-747-3 Sequence 3, Appli
25	69	8.4	406	9	US-10-950-747-3 Sequence 3, Appli
26	69	8.4	406	9	US-10-950-747-3 Sequence 3, Appli
27	69	8.4	406	9	US-10-330-773-602 Sequence 1, Appli
28	69	8.4	406	11	US-11-186-669-1 Sequence 1, Appli
29	69	8.4	406	11	US-11-227-340-1 Sequence 1, Appli
30	69	8.4	444	8	US-10-511-937-2962 Sequence 2962, Ap
31	69	8.4	444	8	US-10-511-937-3014 Sequence 3014, Ap
32	69	8.4	444	10	US-11-183-218-8 Sequence 8, Appli
33	69	8.4	444	11	US-11-183-205-8 Sequence 8, Appli
34	69	8.4	641	11	US-11-227-340-8 Sequence 8, Appli
35	69	8.4	671	11	US-11-029-003-6 Sequence 6, Appli
36	69	8.4	679	11	US-11-227-340-11 Sequence 11, Appli
37	69	8.4	701	11	US-11-227-340-6 Sequence 6, Appli
38	68.5	8.3	462	11	US-11-255-794-12 Sequence 12, Appli
39	68.5	8.3	467	11	US-11-255-794-5 Sequence 5, Appli
40	68.5	8.2	467	11	US-11-036-272-32 Sequence 32, Appli
C 41	68	8.2	427	11	US-11-079-463-6908 Sequence 6908, Ap
C 42	67.5	8.2	285	11	US-11-079-463-8891 Sequence 8891, Ap
43	67.5	8.2	393	11	US-11-188-298-11630 Sequence 11630, A
44	67.5	8.2	477	11	US-11-098-686-10915 Sequence 10915, A
45	67	8.1	269	11	US-11-079-463-8166 Sequence 8166, Ap

ALIGNMENTS

RESULT 1

US-11-072-512-3674

; Sequence 3674, Application US/11072512

; Publication No. US20060029945A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMSECHKA, ICHIRO

; APPLICANT: SEKI, NAOHICO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOKYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: Novel full length cDNA

; FILE REFERENCE: 084335-0191

; CURRENT APPLICATION NUMBER: US/11/072,512

; CURRENT FILING DATE: 2005-03-07

; PRIOR APPLICATION NUMBER: US 60/350,978

; PRIOR FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3674

Alignment Scores:
Pred. No.: 3,79e-89 Length: 158
Score: 785.00 Matches: 150
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 95.0% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-11-072-512-3674 (1-158)

QY 1 ATGGCGTTAGAAAGTCTTGTATGCTCTCGCTGCTTGTGATTGGACCGGTCTGAGAACCTC 60
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAATAAAGTTGCTCTCTGGACTGGTTCATGCTCTCAGTTATCCAGTTTCAGAA 120
Db 21 HisValLysIleSerCysSerLeuAspTrpLeuMetAlaSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
Db 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATGATGATATGATGATATATCTTGTTCGTGATTTGGCATCAGGACA 240
Db 61 ArgIleHisThrTyrlValTyrlGluPheIleTyrlLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGCTAGTTTCTGAGGAACTCTCCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrlPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
QY 361 CTTACACAGTTTCTACTCAGAAATCAATAAATTGGATCCTAGTCCCTTTATTCCTGAC 420
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTTCAGACACAGCAGAGAGTATGATTA 453
Db 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151

RESULT 2

US-10-475-075-788
; Sequence 788, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 788
; LENGTH: 148
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-475-075-788

Alignment Scores:
Pred. No.: 6,52e-88 Length: 148
Score: 775.00 Matches: 148
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 93.8% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x US-10-475-075-788 (1-148)

QY 1 ATGGCGTTAGAAAGTCTTGTATGCTCTCGCTGCTTGTGATTGGACCGGTCTGAGAACCTC 60
Db 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAATAAAGTTGCTCTCTGGACTGGTTCATGCTCTCAGTTATCCAGTTTCAGAA 120
Db 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTCCGGATGAATACATCTGGGAATGGGCTGCCCTGCAAT 180
Db 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATGATGATATGATGATATATCTTGTTCGTGATTTGGCATCAGGACA 240
Db 61 ArgIleHisThrTyrlValTyrlGluPheIleTyrlLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGCTAGTTTCTGAGGAACTCTCCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300
Db 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrlPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
QY 361 CTTACACAGTTTCTACTCAGAAATCAATAAATTGGATCCTAGTCCCTTTATTCCTGAC 420
Db 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTTCAGACACAGCAGAGAGTGA 444
Db 141 PheGlnThrThrAlaGluLeu 148

RESULT 3

US-10-475-075-787
; Sequence 787, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 787
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1

QY	127	AATCTGTATATATTTCGGATGAATATACATCTTGGAAATGGCTGCCCTGCA-----	171
QY	127	AATCTGTATATATTTCGGATGAATATACATCTTGGAAATGGCTGCCCTGCA-----	171
Db	456	-----AlaAspAlaAlaHisLeuGluSerGlyGlnProValGlnCysGlu	470
QY	178	AATCGGATACATCATATGTATATGAGTTTATATATCTTGTTCGTGATTGTGGCATCAGG	237
Db	471	SetAspValIleMetTyrIleGlnGluCysGluGluValIleArgGlnLeuGlnValAsp	490


```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-779

Alignment Scores:
Pred. No.: 0.107
Score: 85.50
Percent Similarity: 41.5%
Best Local Similarity: 25.4%
Query Match: 10.4%
DB: 9
Length: 5415
Matches: 36
Conservative: 23
Mismatch: 58
Indels: 25
Gaps: 5

US-10-664-025-43_COPY_53_511 (1-459) x US-10-995-561-779 (1-5415)

```

QY	/	TTAGAGGCTTTGATGCTCTCTGCTCTGTGATTTGGACCCGGTCTGAGAACTTCATGTG	66
Db		: : : : :	
QY		426 LeuGluLeuLeuLeuGluInAlaAasnLysIleGlnAasnGlyAlaLeuAasnCysGluGlu	445
Db		: : : : :	
QY		67 AAAATAAGTTGCTCTCTGGAAGCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGA	126
Db		: : : : :	
QY		446 LysLeuThrLeuAlaLysAasnThrLeuGln	455
Db		: : : : :	

QY AATCTGTATATATTGCGGAAGAATTTACATCTGGGAATGGGCATGCCCTGCA----- 177

436	-----AlaAspAlaAlaHisLeuGluGluSerGlyGlnProValGlnCysGlu	
D5		
Qy	178 AATCGGATACACATATGTATATGAGTTTATATATCTGTTCGTGATGTGCACAGG	237
	::: ::: ::: ::: ::: :::	
Db	471 SerAspValIleMetTyrIleGlnGluCysGluGlyLeuIleArgGlnLeuGlnValAsp	490
Qy	238 ACAAGGGTAGTTTCTGAGGAACACTCCCTTTTCAACCGAGCTGCTACTCTTACCCCAAGG	297

QY	298	AATATAGATCATGACCCCTCAGGAATCCATTGGAGTGTTCACC---TCTAGGAAA---	351
DB	491	DEUGUUIREUEWASGPURUSNIYYIGUNUEUGUGUULUUAIAIRNEARGVALMET	530
DB	511	ArgLeuGlnAspGluLeuValThrLeuArgLeuGluCysThrAsnLeuTyxArglysGly	530

Db 531 HisPheThrSerLeuGluLeuValProSerThrLeuThrThrHisLeuIysAla 550
QY 397 GATCCT 402
Db 551 GluPro 552

RESULT 10
US-10-995-561-775
; Sequence 775, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michèle et al

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1-001559

```

; CURRENT APPLICATION NUMBER: US107995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 775

```

; LENGTH: 5464
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 US-10-595-561-775

Alignment Scores:	
Pred. No.:	0.107
Score:	85.50
Percent Similarity:	41.5%
Best Local Similarity:	25.4%
Query Match:	10.4%
DB:	9
Length:	5464
Matches:	36
Conservative:	23
Mismatches:	58
Indels:	25
Gaps:	5

US-10-664-025-43_COPY_53_511 (1-459) x US-10-995-561-775 (1-5464)

```
QY 7 TTAGAGCTTCTGATGCTCTCGCTGTCTGATTGGACGGTGTGAGAACCTCCATGTG 66
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 LeuGluLeuLeuGlnIleAlaAsnIlyIleGlnAsnGlyAlaLeuAsnCysGluGlu 445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 67 AAAATAAGTTGCTCTCTGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGCAGA 126
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 LysLeuThrLeuAlaLysAsnThrLeuGln----- 455
QY 127 AATCTGTATATTTCCGGATGAATACATCTGGGAATGGCTGCCCTGCA----- 177
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 456 -----AlaAspAlaAlaHisLeuGluSerGlyGlnProValGlnCysGlu 470
QY 178 AATCGATACATACATATGATATGATTTATATATCTTGTCGTGATGTCGGCATCAGG 237
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 SerAspValIleMetTyrIleGlnGluCysGluGlyLeuIleArgGlnLeuGlnValAsp 490
QY 238 ACAAGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGCTACTTACCCCAAGG 297
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 LeuGlnIleLeuArgAspIleAsnTyrTyrGlnLeuGluGluLeuAlaPheArgValMet 510
QY 298 AATATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACC-----TCTAGGAAA--- 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 ArgLeuGlnAspGluLeuValThrLeuArgLeuGluCysThrAsnLeuTyrArgLysGly 530
QY 352 -----TCAGTGGCTTACACAGTTTCT-----ACTGGAATGAATAAAATTG 396
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 HisPheThrSerLeuGluLeuValProSerThrLeuThrThrHisLeuLysAla 550
QY 397 GATCCT 402
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 GluPro 552
```

RESULT 11

```
US-11-096-568A-8621
; Sequence 8621, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8621
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(242)
; OTHER INFORMATION: Ceres Seq. ID no. 15169117
US-11-096-568A-8621
```

```
Alignment Scores:
Pred. No.: 0.578 Length: 242
Score: 76.50 Matches: 40
Percent Similarity: 37.1% Conservative: 29
Best Local Similarity: 21.5% Mismatches: 70
Query Match: 9.3% Indels: 47
DB: 11 Gaps: 9
```

US-10-664-025-43_COPY_53_511 (1-459) x US-11-096-568A-8621 (1-242)

```
QY 25 CTCGCTGCTTGATTGGACCGTGTGAGAACCTCCAT-----GTGAAATAAGT 75
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 LeuThrLeuAlaLeuTrpPheSerArgAspGlySerHisAspGluAspMetLysLeu 40
QY 76 TGCTCTCGGACTGGTTGATGTC-----TCAGTTATCCAGTTGCA 117
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 41 SerLeuLeuSerGlnHisLeuLeuHisGluAsnValAlaAspSerTyrLeuProLeuPro 60
QY 118 GAAAGCAGAAATCTGTATATATTT-----GGGATGATATTACATCTGGGAATG 165
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AlaSerSerAsnMetTyrTrpPheSerLysAspGlnAlaSerAsnTyrGlnPheGlyPhe 80
QY 166 GCCTGCCCTCGAAATCGGATACATACATATGTATATGATTTATATATCTTGTCTGTGAT 225
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 AsnIleSerTrpAlaArgLeuHisValLeuGlyTyrAsp---IleTyrIleSerGlnAsp 99
QY 226 TGTGCATCAGACACA-----AGGTAGTTTCTGAG 255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100 SerSerCysGluSerAspValPheGluLeuLeuValLysProValArgLeuValArgAsp 119
QY 256 GAAACTCTCCTTTTCAAAACCGAGCTGTACTTTACCCCAAGGAATATAGATCATGACCT 315
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 AsnGluLeuLeuAspGlnGluPheVal-----AsnIleMethHisAlaLeu 134
QY 316 CAGGAAATCCATTTGGAGTGT-----TCCACC 342
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 GlnValValHisPheTyrCysTrpLysGlySerAlaLeuLeuAsnMetSerAsnIle 154
QY 343 TCTAGGAATCAGTGTGGCTTACACAGTTTCTACTGAG-----AATGAA 387
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 AspSerArgValLeuLysValThrGluValGlnArgGluLysIleSerGlyLeuAsnSer 174
QY 388 ATAAATTTGGATCCTAGTCTCCTTTTATTGCTGACTTT-----CAGACACAGCAGAAG 441
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ValLeuLeuAsnAspAsnAspPheAlaSerLysValPheCysAsnMetProSerGluGlu 194
QY 442 TTAGGATTATTAATCTTC 459
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LysGlyCysIleCysPhe 200
```

RESULT 12

```
US-11-096-568A-8620
; Sequence 8620, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8620
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(394)
; OTHER INFORMATION: Ceres Seq. ID no. 15169116
US-11-096-568A-8620
```

```
Alignment Scores:
Pred. No.: 0.664 Length: 394
Score: 76.50 Matches: 40
Percent Similarity: 37.1% Conservative: 29
Best Local Similarity: 21.5% Mismatches: 70
Query Match: 9.3% Indels: 47
DB: 11 Gaps: 9
```

US-10-664-025-43_COPY_53_511 (1-459) x US-11-096-568A-8620 (1-394)

```
QY 25 CTCGCTGCTTGTGATTGGACCGTGTGAGAACCTCCAT-----GTGAAATAAGT 75
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 LeuThrLeuAlaLeuTrpPheSerArgAspGlySerHisAspGluAspMetLysLeu 192
QY 76 TGCTCTCTGAGACTGGTTGATGTC-----TCAGTTATCCAGTTGCA 117
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 SerLeuLeuSerGlnHisLeuLeuHisGluAsnValAlaAspSerTyrLeuProLeuPro 212
```

```
QY 118 GAAAGCAGAAATCTGTATATATT-----GCGGATGAATTACATCTCGGAATG 165
Db 213 AlaserSerAsnMetTyrTrpPheSerLysAspGlnAlaSerAenTyrGlnPheGlyPhe 232
QY 166 GCGTGCCTCGCAATCGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGAT 225
Db 233 AsnileSerTrpAlaArgLeuHisValLeuGlyTyrAsp---ileTyrIleSerGlnAsp 251
QY 226 TGTGGCATCAGGACA-----AGGTAGTTTCTGAG 255
Db 252 SerSerCysGluSerAspValPheGluLeuValLysProValArgLeuValArgAsp 271
QY 256 GAAACTCTCTCTTTTCAACCGAGCTGACTTACCCCAAGGATATAGATGACCCCT 315
Db 272 AsnGluLeuLeuAspGlnGluPheVal-----AsnileMetHisAlaLeu 286
QY 316 CAGGAAATCCATTTGGAGTGT-----TCCACC 342
Db 287 GlnValValHisPheTyrCysTrpLysGlySerAlaLeuLeuAsnMetSerAsnile 306
QY 343 TCTAGGAAATCAGTGTGGCTTACACCGAGTTTCTACTGAG-----AATGAA 387
Db 307 AspSerArgValLeuLysValThrGluValGlnArgGluLysIleSerGlyLeuAsnSer 326
QY 388 ATAAATTTGGATCCTAGTCTCTTTTATTGCTGACTTT-----CAGACACAGCAGAGAG 441
Db 327 ValLeuLeuAsnAspAsnAspPheAlaSerLysValPheCysAsnMetProSerGluGlu 346
QY 442 TTAGGATTATTATCTTC 459
Db 347 LysGlyCysIleCysPhe 352
```

RESULT 13

```
US-11-096-568A-8619
; Sequence 8619, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8619
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(398)
; OTHER INFORMATION: Ceres Seq. ID no. 15169115
US-11-096-568A-8619
```

```
Alignment Scores:
Pred. No.: 0.666 Length: 398
Score: 76.50 Matches: 40
Percent Similarity: 37.1% Conservative: 29
Best Local Similarity: 21.5% Mismatches: 70
Query Match: 9.3% Indels: 47
DB: 11 Gaps: 9
```

US-10-664-025-43_COPY_53_511 (1-459) x US-11-096-568A-8619 (1-398)

```
QY 25 CTCGCTGTCTGATTGGACCGTGTGAGACCTCCAT-----GTGAAATAAGT 75
Db 177 LeuThrLeuAlaLeuTrpPheSerArgAspGlySerHisAspGluAspMetLysLeuIle 196
QY 76 TGTCTCTCTGACTGGTGTGATGTC-----TCAGTTATCCAGTTGCA 117
Db 197 SerLeuLeuSerGlnHisLeuLeuHisGluAsnValAlaAspSerTyrLeuProLeuPro 216
```

```
QY 118 GAAAGCAGAAATCTGTATATATT-----GCGGATGAATTACATCTCGGAATG 165
Db 217 AlaserSerAsnMetTyrTrpPheSerLysAspGlnAlaSerAenTyrGlnPheGlyPhe 236
QY 166 GCGTGCCTCGCAATCGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGAT 225
Db 237 AsnileSerTrpAlaArgLeuHisValLeuGlyTyrAsp---ileTyrIleSerGlnAsp 255
QY 226 TGTGGCATCAGGACA-----AGGTAGTTTCTGAG 255
Db 256 SerSerCysGluSerAspValPheGluLeuValLysProValArgLeuValArgAsp 275
QY 256 GAAACTCTCTCTTTTCAACCGAGCTGACTTACCCCAAGGATATAGATGACCCCT 315
Db 276 AsnGluLeuLeuAspGlnGluPheVal-----AsnileMetHisAlaLeu 290
QY 316 CAGGAAATCCATTTGGAGTGT-----TCCACC 342
Db 291 GlnValValHisPheTyrCysTrpLysGlySerAlaLeuLeuAsnMetSerAsnile 310
QY 343 TCTAGGAAATCAGTGTGGCTTACACCGAGTTTCTACTGAG-----AATGAA 387
Db 311 AspSerArgValLeuLysValThrGluValGlnArgGluLysIleSerGlyLeuAsnSer 330
QY 388 ATAAATTTGGATCCTAGTCTCTTTTATTGCTGACTTT-----CAGACACAGCAGAGAG 441
Db 331 ValLeuLeuAsnAspAsnAspPheAlaSerLysValPheCysAsnMetProSerGluGlu 350
QY 442 TTAGGATTATTATCTTC 459
Db 351 LysGlyCysIleCysPhe 356
```

RESULT 14

```
US-11-188-298-7840
; Sequence 7840, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 7840
; LENGTH: 460
; TYPE: PRT
; ORGANISM: GIBBERELLA ZEAE PH-1
US-11-188-298-7840
```

```
Alignment Scores:
Pred. No.: 1.89 Length: 460
Score: 73.00 Matches: 31
Percent Similarity: 38.2% Conservative: 21
Best Local Similarity: 22.8% Mismatches: 36
Query Match: 8.8% Indels: 48
DB: 11 Gaps: 6
```

US-10-664-025-43_COPY_53_511 (1-459) x US-11-188-298-7840 (1-460)

```
QY 100 TCAGTTATCCAGTTGCAGAAAGCAGAAATCTGTATATATTTTCGGATGAATTA----- 153
Db 115 SerLeuValLysThrAlaHisSerLysAsnIleTyrValMetAlaAspValAlaAsn 134
QY 154 CATCTGGGAATGGCTGCCCTCCAAATCGGATACATATATGTATATAGTTTATATAT 213
Db 135 HisMetGlyLysGlyIleGlnAspHisArgProGluPro----- 147
QY 214 CTTGTTCTGTGATTGGCATCAGGACAGGGTAGTTTCTGAGGAAACTCTCCTTTTCAA 273
Db 148 -----LeuAsnGln 150
```

```
QY 274 ACCGAGCTGTACTTTACCCAGGAATATAGATCATGACCTCTCAGAAATCCATTGGAG 333
Db 151 GlnSerSerTyrHisSerProCysAlaIleAspTyrAsnAsnGlnAsnSerIleGluGln 170
QY 334 TGT-----TCCACCTCTAGGAATCAGTG----- 357
Db 171 CysGluIleAlaAspLeuProAspLeuAsnThrGlySerGluThrValLysLysValLeu 190
QY 358 -----TGCTTACACCGATTTCTACTGAG-----AATGAATAATAATTGGAT--- 399
Db 191 AsnAspTrpIleSerTrpLeuValThrGluTyrSerPheAspGlyIleArgIleAspThr 210
QY 400 -----CCTAGTCCTTTTATTGCTGACTTTTCAGACAAACAGCA 435
Db 211 ValLysHisValGluLysSerPheTrpProAspPheGlnLysAlaAla 226
```

RESULT 15

```
US-11-045-004-81
; Sequence 81, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: CHARBIT, ALAIN
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENIO, ENCARN
; APPLICANT: PABLOS, BETRIZ DE
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
; APPLICANT: HAUF, JORG
; APPLICANT: ROSE, MATTHIAS
; APPLICANT: VOSS, HAMUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; PRIOR FILING DATE: 2000-04-11
```

```

; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-11-045-004-81

Alignment Scores:
Pred. No.: 2,87 Length: 440
Score: 71.50 Matches: 45
Percent Similarity: 40.6% Conservative: 26
Best Local Similarity: 25.7% Mismatches: 60
Query Match: 8.7% Indels: 44
DB: 11 Gaps: 12

US-10-664-025-43_COPY_53_511 (1-459) x US-11-045-004-81 (1-440)
QY 3 GGCCTTAGAAGCTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGCTGCTGAGAACCTCCA 62
Db 199 GlyIleAsnHisLeuHisTrpHisArgValPheAspLysAsp----- 212
QY 63 TGTGAAATAAGTTCCTCTCTGGACTGTTGATGCTCAGTTATCCCA----- 111
Db 213 ---GlyThrGluLeuThrGluLysValIleAspGlyLeu-TyrAlaProAspAlaAsnPr 231
QY 112 -----GTTGCAGAAAGCAGAAATCTG---TATATATTTCGGGATGAATTACATCT 158
Db 231 oGlyLysValValGluAsnIleLysAsnMetArgPheLeuTyrGluGlnValLysHisLe 251
QY 159 GGGAAATG---GGCTGCCCTGC AAAATCGGATACATACATATGTATAT----- 201
Db 251 uLysMetLeuProCysPro-----TyrHisArgTyrTyrTyrMetThrAspAlaMe 268
QY 202 -----GAGTTTATATATCTTGTTCGTGATTTGGCATCAGGCAAGGTAGTTCTGCA 254
Db 268 tLeuGluGluGluLeuAlaSerPheLysAsnGluGlyThrArgGlyGluValValLysLy 288
QY 255 G---GAAACTCTCCTTTTCAACCGAGCTGTACTTTTACCCCAAGGAATATAGATCATGA 311
Db 288 sLeuGluAspSerLeuPhe-----GluLeuTyrLysAspPro-----AsnLeuAspTyrLy 305
QY 312 CCCTCAGGAAATC-----CATTTGGAGTGTTCACCTCTTAGG----- 348
Db 305 sProGluGluLeuSerLysArgGlyAlaHisTyrSerAspAlaAlaCysGluIleI 325
QY 349 -AATCAGTGTGG-----CTTACACCGAGTCTTCTACTGAGATGAATAAAA 392
Db 325 eAsnSerIleTyrAsnAsnLysGlyThrValMetValSerThrArgAsnAsnGlyAl 345
QY 393 ATTTGATCCTAGTCCTCTTTTATTGCTGACTTTTCAGACAAACAGCA 435
Db 345 alleAspAspValProTyrAspSerAlaValGluIleThrSer 359
```

Search completed: May 9, 2006, 10:11:35
Job time : 33.85 secs

QY	112			-----GTTGCAGAA	120
Db	36	GlyLysLeuThrSerSerValGluValGluCysLeuGluAlaGluLeuValValThrVal	55		
QY	121	ACCAAAATCTGTAT-----ATATTGGGATGAATTACATCTGGGAATG	165		
Db	56	SerArgAspLeuPheGlyThrGlyLysLeuIleGlnProGluAspLeuThrLeuGlySer	75		
QY	166	--GGCTGC--CCTGCAATCGGATACATATATGATATATATATATATCTTCTTT	219		
Db	76	GluAsnCysArgProLeuValSerValAlaThrAspValValArgPheLysAlaGlnLeu	95		
QY	220	CCTGATTGGCATCAGACACAGGATGTTCTGAGAACTCTCTCTTTTCAAAACCGAG	279		
Db	96	HisGluCysSerAsnArgValGln--ValThrGluAspAlaLeuValThrSerThrVal	114		
QY	280	CTGTACTTTACCCAGGAATATAGAT-----CATGACCTCAGGAA	321		
Db	115	LeuLeuHisGlnProArgProValProGlyLeuSerIleLeuArgThrAsnArgAlaAsp	134		
QY	322	ATCCATTGGAGTGTCCACCTCTAGG-----AAATCA	354		
Db	135	ValProIleGluCysArgTyrProArgGlnGlyAsnValSerSerHisAlaIleArgPro	154		
QY	355	GTGTGGCTT-----ACACCACTTCTACTGAGAATGAAATA-----	390		
Db	155	ThrTrpValProPheSerThrThrValSerSerGluGluLysLeuValPheSerLeuArg	174		
QY	391	-----AAATTGGATCCTAGTCCTTTTATTGCTGACTTT	423		
Db	175	LeuMetGluGluAsnTrpAsnThrGluLysLeuSerProThrSerHisLeuGlyGluVal	194		
QY	424	CAG-----ACACACAGCAGAGTTAGGATTATTATCTTC	459		
Db	195	AlaTyrLeuGlnAlaGluValGlnThrGlySerHisLeuProLeuLeuLeuPhe	212		
RESULT 2					
T48727					
retinoblastoma binding protein 2 related protein [imported] - Neurospora crassa (fragment)					
N:Alternate names: protein 804.10					
C:Species: Neurospora crassa					
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 31-Dec-2004					
C:Accession: T48727					
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, April 2000					
A:Reference number: Z24541					
A:Accession: T48727					
A:Status: preliminary					
A:Molecule type: DNA					
A:Residues: 1-1009 <SCH>					
A:Cross-references: UNIPROT:Q9C2J9; UNIPARC:UPI000017784E; EMBL:AL353819; GSPDB:GNO0112;					
A:Experimental source: cosmid contig 804; strain 74					
C:Genetics:					
A:Gene: NCSP:804.10					
A:Map position: 2					
A:Introns: 153/2; 480/1					
Alignment Scores:					
Pred. No.:	0.406	Length:	1009		
Score:	86.50	Matches:	47		
Percent Similarity:	32.8%	Conservative:	13		
Best Local Similarity:	25.7%	Mismatches:	45		
Query Match:	10.5%	Indels:	78		
DB:	2	Gaps:	12		
US-10-664-025-43_copy_53_511 (1-459) x T48727 (1-1009)					
QY	340	TGGAACACTCCAAAT-----GGATTTCTCGAGGTCATGAT	305		
Db	493	TrpAsnCysProArgCysLeuValGlyAspGlyGlnPheGlyPheGluGluGly---Gly	501		
QY	304	CTATATTCCTTG-----GGGTAAAGTACAGTCGGTTTGAAGAGGAGAGTTTCCT	254		

Db	502	LeuTyrSerLeuLysGlnPheGlnLysAlaAlaAspPheLysGlnGlyTyrPheGlu	521		
QY	253	CAGAAACTACCTTGTTC-----TGATGCCCAAAAT-----CACGAACAAGATATA	209		
Db	522	AsnLysMetProPheAspAlaValLeuAsnCysHisArgProValThrGluAspAspVal	541		
QY	208	TAAATCTCATATACATATGTATGTATCCGATTTCGAC-----	173		
Db	542	GluArgGluPheTrpArgLeuValAlaAspLeuGluGluThrValGluValGluTyrGly	561		
QY	172	-----GGCAGCCCATTCACGATGTA-----	152		
Db	562	AlaAspIleHisCysThrThrHisGlySerGlyPheProThrValGluLysAsnProAsn	581		
QY	151	-----ATTCATCCGCAAA-----	139		
Db	582	AsnProTyrSerThrAspProTrpAsnLeuAsnLeuLeuProLeuHisPro-GluSerLe	601		
QY	138	-----TATATACAG-----	130		
Db	601	uPheArgTyrIleLysThrAspIleSerGlyMetThrValProTrpValTyrValGlyMe	621		
QY	129	-----ATTTCCTCTTCGCAACTGGGATACTGAGACCATCAACGATCCAGAGGCAAC	74		
Db	621	tIlePheSerThrPheCys---TrpHisAsnGluAspHisTyr---AlaTyrSerAlaAs	639		
QY	73	TTATTTTCCATGCGGTTCTCAGCACCGGT-----CCAAATCAAGACAGCGAGGA	23		
Db	639	nTyrGlnHisPheGlyAlaThrLysThrTyrTyrGlyLeProGlyGluAspAlaGluIy	659		
QY	22	GCATCAA 16			
Db	659	spheGlu 661			
RESULT 3					
T42725					
actin binding protein ACF7, neural isoform 1 - mouse (fragment)					
C:Species: Mus musculus (house mouse)					
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 03-Nov-2000					
C:Accession: T42725					
R:Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.					
Genomics 38, 19-29, 1996					
A:Title: Cloning and characterization of mouse ACF7, a novel member of the dystonin sub					
A:Reference number: Z20900; MUID:97124842; PMID:8954775					
A:Accession: T42725					
A:Status: preliminary; translated from GB/EMBL/DBJ					
A:Molecule type: mRNA					
A:Residues: 1-1825 <BER>					
A:Cross-references: UNIPARC:UPI000016D118; EMBL:U67203; NID:g1675221; PID:g1675222; PIDN					
C:Genetics:					
A:Gene: ACF7					
A:Map position: 4					
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S					
C:Keywords: actin binding					
Alignment Scores:					
Pred. No.:	0.731	Length:	1825		
Score:	84.50	Matches:	36		
Percent Similarity:	41.5%	Conservative:	23		
Best Local Similarity:	25.4%	Mismatches:	58		
Query Match:	10.2%	Indels:	25		
DB:	2	Gaps:	5		
US-10-664-025-43_copy_53_511 (1-459) x T42725 (1-1825)					
QY	7	TTAGAAGTCTTGATGCTCCTCGCTGTCTTGATTTGGACCGGTGCTGAGAACCTCCATGTG	66		
Db	369	LeuGluLeuLeuLeuGlnIleAlaAsnLysIleGlnAsnGlyAlaLeuAsnCysGluGlu	388		
QY	67	AAATAAGTTCTCTCGACTGCTGATGCTTATCCAGTTATCCAGTTTCGAGAAAGCAGA	126		
Db	389	LysLeuThrLeuAlaLysAsnThrLeuGln-----	398		

QY 127 AATCTGTATATATTTCGGATGAATATCATCTCGGAATGGCTGCCCTGCA----- 177
Db 399 -----AlaAspAlaAlaHisLeuGluSerGlyGlnProValGlnCysGlu 413
QY 178 AATCGGATACATACATATGATATGAGTTTATATATCTTTCGTGTGATTTGGCATCAGG 237
Db 414 SerAspValIleMetTyrIleGlnGluCysGluGlyLeuIleArgGlnLeuGlnValAsp 433
QY 238 ACAAGGTAGTTTCTGAGGAACACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGG 297
Db 434 LeuGlnIleLeuArgAspGluLeuLysTyrGlnLeuGluGluLeuAlaPheArgValMet 453
QY 298 AATATAGATCATGACCCCTCAGGAATCCATTTCGAGTGTCCACC---TCTAGGAAA--- 351
Db 454 ArgLeuGlnAspGluLeuValThrLeuArgLeuGluCysThrAsnLeuTyrArgLysGly 473
QY 352 -----TCAGTGTGGCTTACACCAAGTTTCT-----ACTGAGAATGAATAAAATTG 396
Db 474 HisPheSerLeuGluLeuValProProSerThrLeuThrThrThrHisLeuLysAla 493
QY 397 GATCCT 402
Db 494 GluPro 495
RESULT 4
T30847
actin binding protein ACF7, neural isoform 2 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 03-Nov-2000
C;Accession: T30847
R;Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
Genomics 38, 19-29, 1996
A;Title: Cloning and characterization of mouse ACF7, a novel member of the dystronin sub
A;Reference number: Z20900; MUID:97124842; PMID:8954775
A;Accession: T30847
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1885 <BER>
A;Cross-references: UNIPARC:UPI000016D119; EMBL:U67204; NID:g1675223; PID:g1675224; PIDN
C;Genetics:
A;Gene: mACF7
A;Map position: 4
C;Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S

Alignment Scores:
Pred. No.: 0.734 Length: 1885
Score: 84.50 Matches: 36
Percent Similarity: 41.5% Conservative: 23
Best Local Similarity: 25.4% Mismatches: 58
Query Match: 10.2% Indels: 25
DB: 2 Gaps: 5

US-10-664-025-43_COPY_53_511 (1-459) x T30847 (1-1885)

QY 7 TTAGAAGTCTTGATGCTCTCGCTGCTTGAATTTGACCGGTCTCAGAACCTCCATGTG 66
Db 429 LeuGluLeuLeuGlnIleAlaAsnLysIleGlnAsnGlyAlaLeuAsnCysGluGlu 448
QY 67 AAAATAAGTTGCTCTCTGAGCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAACGAGA 126
Db 449 LysLeuThrLeuAlaLysAsnThrLeuGln----- 458
QY 127 AATCTGTATATATTTCGGATGAATTACATCTCGGAATGGCTGCCCTGCA----- 177
Db 459 -----AlaAspAlaAlaHisLeuGluSerGlyGlnProValGlnCysGlu 473
QY 178 AATCGGATACATACATATGATATGAGTTTATATATCTTTCGTGTGATTTGGCATCAGG 237
Db 474 SerAspValIleMetTyrIleGlnGluCysGluGlyLeuIleArgGlnLeuGlnValAsp 493
QY 238 ACAAGGTAGTTTCTGAGGAACACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGG 297

Db 494 LeuGlnIleLeuArgAspGluLysTyrTyrGlnLeuGluGluLeuAlaPheArgValMet 513
QY 298 AATATAGATCATGACCCCTCAGGAATCCATTTCGAGTGTCCACC---TCTAGGAAA--- 351
Db 514 ArgLeuGlnAspGluLeuValThrLeuArgLeuGluCysThrAsnLeuTyrArgLysGly 533
QY 352 -----TCAGTGTGGCTTACACCAAGTTTCT-----ACTGAGAATGAATAAAATTG 396
Db 534 HisPheSerLeuGluLeuValProProSerThrLeuThrThrThrHisLeuLysAla 553
QY 397 GATCCT 402
Db 554 GluPro 555
RESULT 5
T30849
actin binding protein ACF7, neural isoform 3 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30849
R;Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
Genomics 38, 19-29, 1996
A;Title: Cloning and characterization of mouse ACF7, a novel member of the dystronin sub
A;Reference number: Z20900; MUID:97124842; PMID:8954775
A;Accession: T30849
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2033 <BER>
A;Cross-references: UNIPARC:UPI000016D11A; EMBL:U67205; NID:g1675225; PID:g1675226; PID
C;Genetics:
A;Gene: mACF7
A;Map position: 4

Alignment Scores:
Pred. No.: 0.741 Length: 2033
Score: 84.50 Matches: 36
Percent Similarity: 41.5% Conservative: 23
Best Local Similarity: 25.4% Mismatches: 58
Query Match: 10.2% Indels: 25
DB: 2 Gaps: 5

US-10-664-025-43_COPY_53_511 (1-459) x T30849 (1-2033)

QY 7 TTAGAAGTCTTGATGCTCTCGCTGCTTGAATTTGACCGGTCTCAGAACCTCCATGTG 66
Db 577 LeuGluLeuLeuGlnIleAlaAsnLysIleGlnAsnGlyAlaLeuAsnCysGluGlu 596
QY 67 AAAATAAGTTGCTCTCTGAGCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAACGAGA 126
Db 597 LysLeuThrLeuAlaLysAsnThrLeuGln----- 606
QY 127 AATCTGTATATATTTCGGATGAATTACATCTCGGAATGGCTGCCCTGCA----- 177
Db 607 -----AlaAspAlaAlaHisLeuGluSerGlyGlnProValGlnCysGlu 621
QY 178 AATCGGATACATACATATGATATGAGTTTATATATCTTTCGTGTGATTTGGCATCAGG 237
Db 622 SerAspValIleMetTyrIleGlnGluCysGluGlyLeuIleArgGlnLeuGlnValAsp 641
QY 238 ACAAGGTAGTTTCTGAGGAACACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGG 297
Db 642 LeuGlnIleLeuArgAspGluLysTyrTyrGlnLeuGluGluLeuAlaPheArgValMet 661
QY 298 AATATAGATCATGACCCCTCAGGAATCCATTTCGAGTGTCCACC---TCTAGGAAA--- 351
Db 662 ArgLeuGlnAspGluLeuValThrLeuArgLeuGluCysThrAsnLeuTyrArgLysGly 681
QY 352 -----TCAGTGTGGCTTACACCAAGTTTCT-----ACTGAGAATGAATAAAATTG 396
Db 682 HisPheSerLeuGluLeuValProProSerThrLeuThrThrThrHisLeuLysAla 701
QY 397 GATCCT 402
Db 701 -----

Db 702 GluPro 703

RESULT 6

TJ18346

MGC1 protein precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: TJ18346

R:Keeler Jr., C.L.; Hnatow, L.L.; Whetzel, P.L.; Dohms, J.E.

Infect. Immun. 64, 1541-1547, 1996

A:Title: Cloning and characterization of a putative cytoadhesin gene (mgcl) from Mycoplasma gallisepticum

A:Reference number: Z18881; MUID:96201559; PMID:8613358

A:Accession: TJ18346

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1122 <KEE>

A:Cross-references: UNIPROT:Q49379; UNIPARC:UPI00001255FB; EMBL:U34842; NID:g1022701; PMID:8613358

C:Genetics:

A:Gene: mgcl

A:Genetic code: SGC3

Alignment Scores:

Pred. No.:	2.5	Length:	1122
Score:	79.50	Matches:	37
Percent Similarity:	44.6%	Conservative:	29
Best Local Similarity:	25.0%	Mismatches:	55
Query Match:	9.6%	Indels:	27
DB:	2	Gaps:	5

US-10-664-025-43_COPY_53_511 (1-459) x TJ18346 (1-1122)

QY 419 TCAGCAATATAAGGACTAGGATCCAAAT---TTTATTTCATTCTCAGTAGMAACTGGTGTA 363

Db 492 SerSerProArgGlyAlaAsnGlyAsnIlePheLeuGlySerAlaIleThrTrpGlyThr 511

QY 362 AGCCACACTGATTTCTCAGAGTGGA---CACTCAAATGATTTCTCAGAGGTCATGA 306

Db 512 AsnGlyGlyAsnPheLeuAspThrLysTrpHisSerProAlaValIleGluAspAlaPro 531

QY 305 TCTATATTCCTGGGTAAGTACAGCTCGGTGTTGAAAAGAGAGAGTTTCTCAGAAACT 246

Db 532 ThrThrPheIleThrValAsnSerSerGlyValLeuGlnAsnSerGlySerGlnGlnSer 551

QY 245 ACCCTTGTCCTGATGCCAAATCAGCA-----ACAAGATATATAAAC 204

Db 552 ThrSerThrProMetProAsnSerAsnGlyAsnGluSerIleProTyrArgTrpThrAsn 571

QY 203 TCATATACATATGATATGATTCGATTTGCAGGGCAGCCCATTCCTCAGATGTAATTCATCC 144

Db 572 SerTyrAspTyrAsnSerValArgPheAlaAa----- 582

QY 143 GCMAATATATACAGATTTCGTCTTCT-----GCAACTGGATTAACAGACCATCAAC 90

Db 583 -----LeuIleSerLysProAlaGlyGlyAsnThrLysGlnValGlu 596

QY 89 CAGTCCAGAGACAACTTATTTCACATGAGGTTCTCAGCACCGGTCCAAATCAAGACA 30

Db 597 Ser--LeuPheThrThrAlaLeuLysLeuAspThrLeuAsnSerLeuProAsnLysPheTr 616

QY 29 GCGAGGAGCATCAAGACTTC 10

Db 616 hrGlnGluAsnAsnIlePhe 622

RESULT 7

G70519

hypothetical protein Rv1914c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: G70519

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: G70519

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-135 <COL>

A:Cross-references: UNIPROT:Q07719; UNIPARC:UPI00000D1020; GB:Z97193; GB:AL123456; NID:g

C:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1914c

C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1914c

Alignment Scores:

Pred. No.:	2.49	Length:	135
Score:	78.50	Matches:	17
Percent Similarity:	53.1%	Conservative:	9
Best Local Similarity:	34.7%	Mismatches:	14
Query Match:	9.5%	Indels:	9
DB:	2	Gaps:	2

US-10-664-025-43_COPY_53_511 (1-459) x G70519 (1-135)

QY 37 ATTTGGACCGTGTGAGAACCTCATGTGAAATAAGTCTCTCTCGACTGGTTGATG 96

Db 40 LeuTrpValGlyAlaGlySerLeuHisValLys-----MetGlyTrpAlaPhe 55

QY 97 GTCTCAGTTATCCAGTT-----GCAGAAAGCAGAAATCTGTATATATTT 141

Db 56 AlaAlaAspIleProLeuThrSerIleThrLysAlaGluAlaThrAsnAlaArgValTyr 75

QY 142 GCGGATGAATTACATCTCGGAATGGC 168

Db 76 AlaAlaGlyValHisPheGlyPheGly 84

RESULT 8

A24699

stearoyl-CoA 9-desaturase (EC 1.14.19.1) (similarity) - rat

N:Alternate names: acyl-CoA desaturase; Delta9 fatty acid desaturase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004

C:Accession: A24699; JX0150; I58133

R:Thiede, M.A.; Osols, J.; Strittmatter, P.

J. Biol. Chem. 261, 13230-13235, 1986

A:Title: Construction and sequence of cDNA for rat liver stearyl coenzyme a desaturase.

A:Reference number: A24699; MUID:87008535; PMID:2428815

A:Accession: A24699

A:Molecule type: mRNA

A:Residues: 1-358 <THI>

A:Cross-references: UNIPROT:P07308; UNIPARC:UPI000004210A; GB:J02585; NID:g206859; PMID:

R:Mihara, K.

J. Biochem. 108, 1022-1029, 1990

A:Title: Structure and regulation of rat liver microsomal stearoyl-CoA desaturase gene.

A:Reference number: JX0150; MUID:91210202; PMID:1982442

A:Accession: JX0150

A:Molecule type: DNA

A:Residues: 1-290, 'A', 292-358 <MIH>

A:Cross-references: UNIPARC:UPI00000E69BF

R:Baba, H.; Fuss, B.; Watson, J.B.; Zane, L.T.; Macklin, W.B.

Neurochem. Res. 19, 1091-1099, 1994

A:Title: Identification of novel mRNAs expressed in oligodendrocytes.

A:Reference number: I58133; MUID:95098214; PMID:7800118

A:Accession: I58133

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-4, 'I', 6-10, 'G', 12-13, 'SA', 16-20, 'A', 22-25, 'GQ', 28-30, 'G', 32-33, 'PE', 36, 'N',

A:Cross-references: UNIPARC:UPI00000E198; GB:S75730; NID:g861452; PMID:AAB32826.1; PID:

C:Superfamily: acyl-CoA desaturase; stearoyl-CoA desaturase homology

C:Keywords: oxidoreductase; unsaturated fatty acid biosynthesis

F:113-307/Domain: stearoyl-CoA desaturase homology <SDH>

Alignment Scores:

Pred. No.:	4.13	Length:	358
------------	------	---------	-----

Score: 77.00 Matches: 38
Percent Similarity: 39.5% Conservative: 22
Best Local Similarity: 25.0% Mismatches: 50
Query Match: 9.3% Indels: 42
DB: 1 Gaps: 8

US-10-664-025-43_COPY_53_511 (1-459) x A24699 (1-358)

QY 124 AGAAATCTGTATATATTGGCGAATGAAATACATCTGGGA-----ATGGGC 168

DB 73 ArgasnlleleuMetAla---LeuLeuHisValGlyAlaLeuTyGlyIleThrLeu 91

QY 169 TGGCCTGCAATCGGATACATACATATGATGAG---TTTATATATCTGTTCGGTAT 225

DB 92 IleProSerSerLysValTyThrLeuLeuTrpGlyIlePheTyThrLeuIleSerAla 111

QY 226 TGTGGCATC----- 234

DB 112 LeuGlyIleThrAlaGlyAlaHisArgLeuTrpSerHisArgThrTyThrLysAlaArgLeu 131

QY 235 -----AGGACAGGGTAGTTCTGAGGAACCTCTCTTTTCAACCGAGCTGTACTTT 288

DB 132 ProLeuargIlePheLeuIleAlaAsnThrMetAlaPheGlnAsnAspValTyGlu 151

QY 289 ACCCCAGGAAT-----ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTCCACC 342

DB 152 TrpAlaArgAspHisArgAlaHisLysPheSerGluThrHisAlaAspProHisAsn 171

QY 343 TCTAGGAATCAGTG-----TGGCTTACA-----CCAGTT 372

DB 172 SerArgArgIlePhePheSerHisValGlyTrpLeuLeuValArgLysHisProAla 191

QY 373 TCTACTGAGATGAATAAATTGGATCCTAGTCCCTTTTATTGCTGAC----- 420

DB 192 ValLysGluLysGlyLysLeuAspMetSerAspLeuLysAlaGluLysLeuValMet 211

QY 421 TTTCCAGACACAGCAGAGAGTAGGATTTAATATC 456

DB 212 PheGlnArgTyThrLysProGlyLeuLeuLeu 223

RESULT 9

A82094

thiamin biosynthesis lipoprotein ApbE VC2289 [imported] - Vibrio cholerae (strain N16961

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: A82094

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: A82094

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-367 <HEI>

A:Cross-references: UNIPROT:Q9KPS3; UNIPARC:UPI000000C3262; GB:AE004300; GB:AE003852; NID

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2289

A:Map position: 1

C:Superfamily: hypothetical protein HI0172

Alignment Scores:

Pred. No.: 4.71 Length: 367

Score: 76.50 Matches: 39

Percent Similarity: 41.4% Conservative: 21

Best Local Similarity: 26.9% Mismatches: 42

Query Match: 9.3% Indels: 43

DB: 2 Gaps: 9

US-10-664-025-43_COPY_53_511 (1-459) x A82094 (1-367)

QY 55 AACCTCCATCTGAAATAAGTTGCTCTCTGAGCTGGTTGATGGTCTCAGTTATCCAGTT 114

DB 5 AsnTrpIleLeuLysValSerMetValIle---TrpPheArgValSerValPheProIle 23

QY 115 GCA-----GAAAGCAGAAATCTGTATATATTGCGGATGAA 150

DB 24 SerIlePheAsnTyrgluArgSerLysGlnValArgAsnTrpLeuValAlaLeuAlaSer 43

QY 151 TTATCATCTGGGAATGGGCTGC-----CTGCAAAATCGGATACAT----- 189

DB 44 LeuLeuLeuLeuAlaGlyCysGluLysProAlaGluGlnValHisLeuSerGlyProThr 63

QY 190 -----ACATATGTATATGATTTATATATCTTCTCGTGATTTGGCATCAGACA 240

DB 64 MetGlyThrThrTyraenilleLysTyriIleGln-----GlnPro 76

QY 241 AGGCTAGTTTCTCAGGAAACTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 300

DB 77 GlyIleAlaAspSerLysThrLeu-----GlnThrGlu----- 87

QY 301 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 360

DB 88 IleAspArgLeuLeuGluGluValAsnAspGlnMetSerThrTyraGlyLysAspSerGlu 107

QY 361 CTTACACCATGTTCT-----ACTGAGATGAATAAAATGGATCCTAGTCTCTTTATT 414

DB 108 LeuSerArgPheAsnGlnHisThrSerSerGlu-----ProPheAla 121

QY 415 GCTGACTTTTCAGACA 429

DB 122 ValSerThrGlnThr 126

RESULT 10

AE0152

probable bacteriophage protein YP01246 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AE0152

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AE0152

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 <KUR>

A:Cross-references: UNIPROT:Q8ZGP1; UNIPARC:UPI000000DCD86; GB:AL590842; PIDN:CAC90080.1

C:Genetics:

A:Gene: YP01246

Alignment Scores:

Pred. No.: 4.86 Length: 468

Score: 76.50 Matches: 31

Percent Similarity: 33.6% Conservative: 20

Best Local Similarity: 20.4% Mismatches: 31

Query Match: 9.3% Indels: 3

DB: 2 Gaps: 3

US-10-664-025-43_COPY_53_511 (1-459) x AE0152 (1-468)

QY 40 TGGACCGGTGCTGAGAACCTCCATCTGTGAAATAAGTTGCTCTGAGCTGGTTGATGGTC 99

DB 19 TrpGlnTrpSerGluHisLeuHisArgAlaSerPheArgGlyValProPheValVal 38

QY 100 TCAGTTATCCCAAGTTGCAGAAAGCAGAAATCTGTATATATTGCGGATGAATACATCTG 159

DB 39 Ser----- 39

QY 160 GGAATGGGCTGCCTGCAGAAATCGG-----ATACATACATATGTATATGAGTTTATAT 213

DB 40 GlyGlnGlyThrPheGlyArgArgGlnValThrHisSerTyProTyArgAspThrSer 59


```

C:Accession: T33152
R:Davidson, S.; Wohldmann, P.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid T04DL1.
A:Reference number: Z21292
A:Accession: T33152
A:Status: preliminary; translated from GB/EMBL/DDBBJ
A:Molecule type: DNA
A:Residues: 1-2957 <DAV>
A:Cross-references: UNIPROT:O61845; UNIPARC:UPT000007A573; EMBL:AF067617; PIDN:
A:Experimental source: strain Bristol N2; clone T04DL1
C:Genetics:
A:Gene: CESP:T04DL.4
A:Map position: 1
A:Introns: 122/3; 293/3; 515/3; 1205/2; 1577/1; 2221/3; 2776/1; 2864/3

Alignment Scores:
Pred. No.: 6.93
Score: 76.00
Percent Similarity: 43.0%
Best Local Similarity: 24.7%
Query Match: 9.2%
DB: 2

Length: 2957
Matches: 23
Conservative: 17
Mismatches: 35
Indels: 18
Gaps: 3

US-10-664-025-43_COPY_53_511 (1-459) x T33152 (1-2957)

Qy 184 ATACATACATGATATAT-----GAGTTTATA 210
:::|||||
Db 2056 ValHisAsnProValTyrAsnLeuLeuAsnTrnThrTrpGluGluGluPheLeu 2075

```

Qy	211	TATCTTTGTCGATTTGTGGCATCAGGACAAAGGTA-----GTTTCTGAGGAAC	261
		:::	:::
Db	2076	ArgValValArgSerTyrglyVallyThrAsnilecglgluclyGlyAsnGluAspTrp	2095
Qy	262	CTCCTTTTTCACACCGAGCTGTACTTTTACCCCAAGAAATATAGATCATGACCCCTCAGGAA	321
		:::	:::
Db	2096	ThrIlePheIysThrLeuSerProileLeuGluLysLysThrAspGluAspCysHisGlu	2115
Qy	322	-----ATCCATTTGGAGTGTCACCACTCTAGGAAATTCAGTGTGGCTTACACCA	369
		:::	:::
Db	2116	GlnLeuMetCysValleualametCysGlnArgAlaGlnGlySerAsnAspLeuLysPro	2135
Qy	370	GTTCCTACTGAGAAATGAAATATAAATTTGGATCTCTAGTCCT	408
		:::	:::
Db	2136	IleAspLeuLysArgAlaMetSerIleAspProMetPro	2148
RESULT 15			
S70399			
zona pellucida glycoprotein C - cat			
C:Species: Felis silvestris catus (domestic cat)			
C:Date: 28-Oct-1996 #sequence_revision 27-feb-1997 #text_change 09-Jul-2004			
C:Accession: S70399			
R:Harris, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, DNA Seq. 4, 361-393, 1994			
A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a			
A:Reference number: S70396; MUID:95143578; PMID:7841460			
A:Accession: S70399			
A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-424 <HAR>			
A:Cross-references: UNIPROT:P48832; UNIPARC:UPI000013C463; EMBL:U05778; NID:9451			
C:Superfamily: sperm-binding glycoprotein P23; ZP domain homology			
F:43-299/domain: ZP domain homology <ZPH>			
Alignment Scores:			
Pred. No.:		6.21	424
Score:		75.50	42

```

Percent Similarity: 37.1%
Best Local Similarity: 22.6%
Query Match: 9.1%
DB: 2
Conservative: 27
Mismatch: 50
Indels: 67
Gaps: 11
US-10-664-025-43 COPY 53 511 (1-459) x S70399 (1-424)

```

```
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTGCT----- 51
Db ||| ||| : : : ||| : : : ||| : : : |||
1 MetGlyLeuSerTyrGlyLeuPheIleCysPheLeuLeuTrpAlaGlyThrClyLeuCys 20
QY 52 -----GAGAACCTCCAT-----GTG 66
Db : : : |||
21 TyrProThrThrThrGluAspLysThrHisProSerLeuProSerSerProSerVal 40
QY 67 AAATAAGTCTCTCTCGAGCTGGTGGTGGTGGTCTCAGTTATCCCAAGTTGCAGAAAGCAGA 126
Db : : : ||| : : : ||| : : : ||| : : : |||
41 ValValGluCysArgHisAlaTrpLeuValValAsnVal-----SerLys 55
QY 127 AATCTGTAT-----ATATTTGGGATGAATTACATCTGGGAATGGCTGC 171
Db ||| ||| : : : ||| : : : |||
56 AsnLeuPheGlyThrGlyArgLeuValArgProAlaAspLeuThrLeuGly----- 72
QY 172 CTGCAAAAT-----CGGATACATACA 192
Db ||| |||
73 ProGluAsnCysGluProLeuIleSerGlyAspSerAspThrValArgPheGluVal 92
QY 193 TATGTATATGAGTTTATATATCTTCTGCTGATTTGGCATCAGGACAGGTAGTTTCT 252
Db : : : : : ||| ||| : : : ||| : : : |||
93 GluLeuHisLys-----CysGlyAsnSerValGln---ValThr 104
QY 253 GAGGAACTCTCTTTTCAACCGAGCTGTACTTTTACCCCAAGGAATATAGATCAT--- 309
Db ||| : : : ||| ||| ||| ||| : : : |||
105 GluAspAlaLeuValTyrSerThrPheLeuLeuHisAsnProArgProMetGlyAsnLeu 124
QY 310 -----GACCCTCAGGAAATCCATTTTGGAGTGTTCACCTCTAGGAAATCA 354
Db : : : ||| : : : ||| : : : |||
125 SerIleLeuArgThrAsnArgAlaGluValProIleGluCysArgTyrProArgHisSer 144
QY 355 GTGCGCTTACACGAGTTTCTACTGAGAATGAAATAAAATTGGATCCTAGTCTTTTATT 414
Db ||| ||| : : : ||| : : : |||
145 -----AsnValSerSerGluAlaIleLeu-----ProThrTrp 155
QY 415 GCTGACTTTTCAGACAACA 432
Db ||| : : : |||
156 ValProPheArgThrThr 161
```

Search completed: May 9, 2006, 10:10:11
Job time : 25.6 secs

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 9, 2006, 10:02:52 ; Search time 23.58 Seconds
(without alignments)
4120.070 Million cell updates/sec

Title: US-10-664-025-43_COPY_53_511
Perfect score: 826
Sequence: 1 atggcgcttagaagcttctgat.....agttagattattatacttc 459

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2164443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame*_n2p.model -DRV=xlh
-Q=/abses/ABSSWEB_spool/US10664025/runat_08052006_173450_28308/app_query.fasta_1
-DB=Uniprot -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=dct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs07
-USER=US10664025 @CGN_1_1_266 @runat_08052006_173450_28308 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES										
Result No.	Score	Query Match	Length	ID	Description					
1	789	95.5	158	2	Q86WS3 HUMAN					
2	785	95.0	158	2	Q8N9U6 HUMAN					
3	431.5	52.2	164	2	Q4FZG8 MOUSE					
4	234	28.3	173	2	Q9J183 MOUSE					
5	230	27.8	173	2	Q80RW3 MOUSE					
6	223.5	27.1	212	2	Q9BHJ0 HUMAN					
7	219.5	26.6	173	2	Q4V7E2 RAT					
8	142	17.2	202	2	Q925U0 MOUSE					
9	103	12.5	92	2	Q8K4N9 MOUSE					
10	92.5	11.2	422	1	ZP3 MSAU					
11	86.5	10.5	408	2	Q4T0A6 TETNG					
12	86.5	10.5	1736	2	Q9C2J9 NEUCR					
13	85.5	10.4	1062	2	Q6ZSD7 HUMAN					
14	85.5	10.4	5430	1	MACF1 HUMAN					
15	85.5	10.4	5430	2	Q5VM20 HUMAN					
16	85	10.3	1003	2	Q95YL0 CTOSA					

17	85	10.3	1241	2	Q81GP9 BACCR	Q81GP9 bacillus ce
18	85	10.3	1241	2	Q6HM43_BACHK	Q6hm43 bacillus th
19	85	10.3	1241	2	Q63EM2_BACCZ	Q63em2 bacillus ce
20	84.5	10.2	1900	2	Q5DTZ2_MOUSE	Q5dtz2 mus musculu
21	84.5	10.2	5327	1	MACF1_MOUSE	Q9qx0 mus musculu
22	84.5	10.2	7354	2	Q4PL15_MOUSE	Q4pl15 mus musculu
23	84	10.2	526	2	Q6FLR6_CANGA	Q6flr6 candida gla
24	84	10.2	1241	2	Q73C23_BACCL	Q73c23 bacillus ce
25	84	10.2	1241	2	Q81TW1_BACAN	Q81tw1 bacillus an
26	82.5	10.0	272	2	Q8QLN1_NPVMC	Q8qln1 mamestra co
27	82.5	10.0	272	2	Q71AK8_9NUCL	Q71ak8 mamestra co
28	82.5	10.0	486	2	Q804D8_BUFAR	Q804d8 bufo aranan
29	82	9.9	442	2	Q72B14_DESVH	Q72b14 desulfovibr
30	80.5	9.7	2502	2	Q4RQ55_TETNG	Q4rq55 tetraodon n
31	79.5	9.6	418	2	Q8K405_GRODE	Q8k405 lagurus lag
32	79.5	9.6	433	2	Q91L91_WSSV	Q91l91 white spot
33	79.5	9.6	964	2	Q86075_MYCGA	Q86075 mycoplasma
34	79.5	9.6	1122	1	ADP1_MYCGA	Q49379 mycoplasma
35	79	9.6	1241	2	Q4MPH4_BACCE	Q4mph4 bacillus ce
36	79	9.6	2223	2	Q5CYH3_CRYPV	Q5cyh3 cryptospori
37	78.5	9.5	135	2	Q7TZA9_MYCBO	Q7tza9 mycobacteri
38	78.5	9.5	135	2	Q07719_MYCTU	Q07719 mycobacteri
39	78.5	9.5	361	2	Q4T2W2_TETNG	Q4t2w2 tetraodon n
40	78.5	9.5	477	2	Q716P6_9CAUD	Q716p6 mycobacteri
41	78	9.4	284	2	Q948P9_HEVBR	Q948p9 hevea brasi
42	78	9.4	501	2	Q9BID0_ANOGA	Q9bid0 anopheles g
43	78	9.4	501	2	Q7PEP3_ANOGA	Q7pep3 anopheles g
44	77.5	9.4	272	2	Q8JME9_9NUCL	Q8jme9 mamestra co
45	77.5	9.4	511	2	Q6D2D1_ERWCT	Q6d2d1 erwinia car

ALIGNMENTS

RESULT 1
Q86WS3 HUMAN
ID Q86WS3 HUMAN PRELIMINARY; PRT; 158 AA.
AC Q86WS3
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein FLJ36198.
GN Name=FLJ36198;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Buttrick A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maizumi M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.

```
RC TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048121; RAH48121.1; -; mRNA.
DR EMBL; BC036256; RAH36256.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17971 MW; EEB43D6FB8AB81FF CRC64;

Alignment Scores:
Pred. No.: 3,738-82 Length: 158
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 95.5% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x Q8N9S3_HUMAN (1-158)
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTprThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAAATAAGTTGCTCTCTGACGTGGTGTGATGGTCTCAGTTATCCAGTTGAGAA 120
DB 21 HisValValSerGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn 100
QY 121 AGCAGAAATCTGTATATATTTGCGGATGAATTCATCTGGGAATGGGCTGCCCTGCAAT 180
DB 41 SerArgAsnLeuTyPhePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATATATATATCTTGTTCGTGATTTGGCATCAGGACA 240
DB 61 ArgIleHisThrTyPheValTyGluPheIleTyLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGGTAGTTTCTGAGAAACTCTCTCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 300
DB 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTpr 120
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAATTTGGATCCTCTAGTCTTTTATTGCTGAC 420
DB 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTTCAGACACAGCAGAGAGTTAGATTAATTA 453
DB 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

RESULT 2
Q8N9U6_HUMAN PRELIMINARY; PRT; 158 AA.
AC Q8N9U6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ36198.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX PubMed=14702039; DOI=10.1038/ng1285;

Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
Nomura Y., Togiyi S., Komai F., Hara R., Takeuchi K., Arita M.,
Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
DR Nat. Genet. 36:40-45 (2004) .
RL EMBL; AK093157; BAC04191.1; -; mRNA.
DR Ensembl; ENSG00000149507; Homo sapiens.
SQ SEQUENCE 158 AA; 17942 MW; EEB43D70A8A391E8 CRC64;

Alignment Scores:
Pred. No.: 1,098-81 Length: 158
Score: 785.00 Matches: 150
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 95.0% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-43_COPY_53_511 (1-459) x Q8N9U6_HUMAN (1-158)
QY 1 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGACCGGTGCTGAGAACCTC 60
DB 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTprThrGlyAlaGluAsnLeu 20
QY 61 CATGTGAAAATAAGTTGCTCTCTGACGTGGTGTGATGGTCTCAGTTATCCAGTTGAGAA 120
DB 21 HisValValSerGluThrLeuLeuPheGlnThrGluLeuTyPheThrProAlaGlu 40
QY 121 AGCAGAAATCTGTATATATTTGCGGATGAATTCATCTGGGAATGGGCTGCCCTGCAAT 180
DB 41 SerArgAsnLeuTyPhePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
QY 181 CGGATACATACATATGATATATATATATATCTTGTTCGTGATTTGGCATCAGGACA 240
DB 61 ArgIleHisThrTyPheValTyGluPheIleTyLeuValArgAspCysGlyIleArgThr 80
QY 241 AGGGTAGTTTCTGAGAAACTCTCTCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 300
DB 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn 100
QY 301 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
DB 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTpr 120
QY 361 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAATTTGGATCCTCTAGTCTTTTATTGCTGAC 420
DB 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
QY 421 TTTTCAGACACAGCAGAGAGTTAGATTAATTA 453
DB 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
```



```

Qy 421 TTTCAGACACAGCAGAGAGCTTAGGATTATTA 453
Db 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151

RESULT 3
Q4FZG8 MOUSE PRELIMINARY; PRT; 164 AA.
AC Q4FZG8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshitsuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC09498; AAh99498.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 18849 MW; 49ADE19216BE8606 CRC64;

Alignment Scores:
Pred. No.: 1.1e-40 Length: 164
Score: 431.50 Matches: 85
Percent Similarity: 72.4% Conservative: 25
Best Local Similarity: 55.9% Mismatches: 41
Query Match: 52.2% Indels: 1
DB: Gaps: 1

US-10-664-025-43_COPY_53_511 (1-459) x Q4FZG8_MOUSE (1-164)

Qy 1 ATGGCGTTAGAGCTTTGATGCTCCCTGCTGCTTGAATTTGGACCGGCTGAGACCTC 60
Db 1 MetAlaLeuGluValLeuValrLeuAlaValLeuValrThrCysAlaTrpAspIle 20
Qy 61 CATGTGAAATAGTTGCTCTCTGGACTGGTGTCTCAGTTATCCAGTTGCAGAA 120
Db 21 AspValAspValSerCysSerGlnAspTrpMetThrValSerValSerAlaPheSerGln 40
Qy 121 AGC---AGAAATCTGATATATTTCGGATGAATTCATCTGGGAATGGCTGCCCTGCA 177
Db 41 AsnLysArgAsnProTrpIlePheAlaAspGluLeuTyrrLeuGlyGlnAsnCysArgVal 60
Qy 178 AATCGGATACATACATATGATATAGTTATATATCTTGTTCGTGATGTGCATCAGG 237

```

```

Db 61 ThrGlnIleHisAlaHisGlnTyrAspPheIleTyrProValSerHisCysGlyIleArg 80
Qy 238 ACAAGGCTAGTTTCTGAGGAAACTCTCTCTTTTCAAAACCGAGCTGTACTTTACCCCAAGG 297
Db 81 ThrLysValIleSerAsnGluIleValCysPheGluThrGluMetTyrPheArgProArg 100
Qy 298 AATATAGATCATGACCCCTCAGGAAATCCATTTCGAGTGTTCACCTCTAGAAATCAGTG 357
Db 101 AsnTyrCysLeuGluGlnIleValProLeuGlnCysSerAlaSerArgLysSerVal 120
Qy 358 TGCCTTACACCACTTTCTACTGAGAAATGAAATGGAATTTGGATCTCTAGCTCTTTTATGCT 417
Db 121 TrpLeuMetProLeuSerThrGluGluAspProLysProValLysSerProPheMetThr 140
Qy 418 GACTTTCAGACACAGCAGGAGAGTAGGATTATTA 453
Db 141 AspPheGluAlaThrProGluGluLeuGlyLeuLeu 152

RESULT 4
Q9J183 MOUSE PRELIMINARY; PRT; 173 AA.
AC Q9J183;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE EPCS26 (PLAC1) (Mus musculus adult female placenta cDNA, RIKEN full-length enriched library, clone:1600019P04 product:placental specific protein 1, full insert sequence).
GN Names-Plac1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hemberger M.C., Hamelbauer H., Ruschmann J., Zeitz C., Fundele R.H.;
RT "cDNA subtraction cloning reveals novel genes whose temporal and
RT spatial expression indicates association with trophoblast invasion."
RL Dev. Biol. 0:0-0(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20453190; PubMed=10995572; DOI=10.1006/geno.2000.6302;
RA Cocchia M., Huber R., Pantano S., Chen E.Y., Ma P., Forabosco A.,
RA Ko M.S., Schlesinger D.;
RT "PLAC1, an Xq26 gene with placenta-specific expression."
RL Genomics 68:305-312(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

```


RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RC Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051666; AAH51666.1; -; mRNA.
DR MGI; MGI:1926287; Plac1.
DR GO; GO:0005615; C:extracellular space; TAS.
SQ SEQUENCE 173 AA; 19573 MW; 79DDC0179BD730B0 CRC64;

Alignment Scores:
Pred. No.: 2,64e-17 Length: 173
Score: 230.00 Matches: 42
Percent Similarity: 60.7% Conservative: 32
Best Local Similarity: 34.4% Mismatches: 46
Query Match: 27.8% Indels: 2
DB: 2 Gaps: 2

US-10-664-025-43_COPY_53_511 (1-459) x Q80WV3_MOUSE (1-173)
QY 19 ATGCTCTCGCTCTTGTATTTGGACGGTGTGAGAACCTCCATGCGAAATAAGTTCG 78
Db ::
11 ValLeuValAlaPheMetLeuPheSerTyrSerGluGlnAsnGlnValAsnValLeuCys 30
QY 79 TCCTCGAGTGGTGTGAGTCTCAGTTATCCCAAGTTCGAGAAAGCAATCTGTATATA 138
Db ::
31 SerThrAspTrpPheMetValThrValHisProPheLeuAsnAspValTyrVal 50
QY 139 TTTCGGAGCAATTACATCTGGGAATGGCTGCGCTCCCTGCAATCGGATACATATGTA 198
Db ::
51 HisPheTyrGluValHisLeuGlyLeuGlyCysProProAsnHisValHisProArgPhe 70
QY 199 TATGAGTTTATATATCTTGTCTGATTTGGCATCAGCAAGGAGTGTCTGAGGAA 258
Db ::
71 TyrGlnPheHisTyrArgValThrGluCysGlyIleArgIleValSerProAsp 90
QY 259 ACTCTCTTTTCAACCGAGTGTACTTTTACCCCAAGGAATATAGATCATGCCCTCAG 318
Db ::
91 ValValIleTyrSerSerGluIleHisTyrAlaSerLysGly---SerSerThrIleTyr 109
QY 319 GAAATCCATTGAGGTGTTCCACCTTAGGAATAGTGTGCTTACA---CCAGTTTCT 375
Db ::
110 ValIleProValSerCysAlaAlaProArgSerProTrpLeuThrIlePheTyrSer 129
QY 376 ACTGAG 381
Db :::
130 AlaLys 131

RESULT 6
QHBJO HUMAN
ID Q9HBJO HUMAN PRELIMINARY; PRT; 212 AA.
AC Q9HBJO
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE PLAC1 (Hypothetical protein FLJ90605) (Placenta-specific 1).
GN Name=PLAC1.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20453190; PubMed=1095572; DOI=10.1006/geno.2000.6302;
RA Cocchia M., Huber R., Pantano S., Chen E.Y., Ma P., Forabosco A.,
RA Ko M.S., Schlessinger D.;
RT "PLAC1, an Xq26 gene with placenta-specific expression.";
RL Genomics 68:305-312(2000).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Placenta;
RC Otsuki T., Ota T., Nishikawa T., Hayaishi K., Suzuki Y., Yamamoto J.,
RA Wakamatsu A., Kimura K., Sakamoto K., Hatano N., Kawai Y., Ishii S.,
RA Saito K., Kojima S., Sugiyama T., Ono T., Okano K., Yoshikawa Y.,
RA Aotsuka S., Sasaki N., Hattori A., Okumura K., Nagai K., Sugano S.,
RA Tsogai T.;
RT "Signal Sequence and Keyword Trap in silico for Selection of Full-
length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-
capped cDNA Libraries.";
RL DNA Res. 12:117-126(2005).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Placenta;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Placenta;
RC Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[5]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Placenta;
RG NIH MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF234654; AAG22596.1; -; mRNA.
DR EMBL; AK075086; BAC11392.1; -; mRNA.
DR EMBL; BC022335; AAH22335.1; -; mRNA.
DR EMBL; BC066327; AAH66327.1; -; mRNA.
DR Ensembl; ENSG00000170965; Homo sapiens.
DR HGNC; HGNC:9044; PLAC1.
SQ SEQUENCE 212 AA; 23616 MW; D449300A597DD023 CRC64;

Alignment Scores:
Pred. No.: 1.58e-16 Length: 212
Score: 223.50 Matches: 40
Percent Similarity: 60.0% Conservative: 32
Best Local Similarity: 33.3% Mismatches: 43
Query Match: 27.1% Indels: 5
DB: 2 Gaps: 2

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nkaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Scaubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bareh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nkaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Varardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakayama N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume M., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [6]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.152600;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukaya S., Furuno M., Hanagaki T., Hara A., Hashizume M.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB050982; BAB47553.1; -; mRNA.
DR ENBL; AF420487; AAL82591.1; -; mRNA.
DR ENBL; AK033038; BAC28135.1; -; mRNA.
DR Ensembl; ENSMUSG00000041857; Mus musculus.
DR MGI; MGI:2149290; Oospl.
DR GO; GO:0005615; C:extracellular space; TAS.
FT SIGNAL.
SQ SEQUENCE 202 AA; 23013 MW; 6CD040BF26BB95FB CRC64;
Alignment Scores: 4.42e-07 Length: 202
Pred. No.: 142.00 Matches: 41
Percent Similarity: 45.8% Conservative: 29
Best Local Similarity: 26.8% Mismatches: 65
Query Match: 17.2% Indels: 18
DB: 2 Gaps: 7

US-10-664-025-43_COPY_53_511 (1-459) x Q925U0_MOUSE (1-202)

QY 7 TTAGAAGTCTTGATGCTCTCGTGTCTG-----ATTGGACCGGTGCT 51
Db 8 LeuGlyLeuLeuLeuLeuSerPheMetLysThrCysAlaAspAspTrpThrAla--- 26
QY 52 GAGAACCTCCATGTGAAATAAGTTGCTCTCTGGACTGGTGGTATGCTCAGTTATCCCA 111
Db 27 -----IleSerLeuGlnCysAlaAspHisTrpPheHisLeuArgIleArgPro 42
QY 112 GTTGCAAGAGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGC 171
Db 43 ThrIle---PheHisAsnIlePheMetGluProAspGluValPheLeuGlyIleGlyCys 61
QY 172 CTGCGAAATCGGATACATACATAT-----GTATATGATGTTTATATATCTTGTTCGT 222
Db 62 ProValThr-----ThrTrpProAsnAspThrTrpGluPheIleTrpArgThrTyr 79
QY 223 GATTGTGGCATCAGGACAGAGGTAGTTTCTGAGGAAATCTCTCTTTTCAACCCGAGCTG 282
Db 80 SerCysGlyIleAlaAsnLysValLeuCysAspValThrLeuLeu---LysThrGlnLeu 98
QY 283 TACTTTACCCCAAGGAATATAGATCATGACCTTCAGGAAATCCATTTGGAGTGTCCACC 342
Db 99 ThrTrpIleSerLysAsnAlaSerLeuGln---AlaGluMetSerLeuSerCysValMet 117
QY 343 TCTAGGAATCAGTGGCTTACACCAGTTTCTACTGAGATGAATAAATTTGGATCCT 402
Db 118 HisAsnGlnSerProHisPheCysGluAlaGluSerArgGlyAspPheThrGlyAspPro 137
QY 403 AGTCCTTTTATTCGTCATCTTCAGTTCACACACACAGCAGAGAG 441
Db 138 ProGlyTrpThrGluAspMetArgAlaArgArgAspGlu 150

```
RESULT 9
Q8K4N9 MOUSE
ID Q8K4N9 MOUSE PRELIMINARY; PRT; 92 AA.
AC Q8K4N9.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Initiate factor 3.2.
GN Name=Oosp1; Synonyms=IF3 2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=2222242; PubMed=12237121; DOI=10.1016/S0006-291X(02)02194-0;
RX Mano H., Nakatani S., Aoyagi R., Ishii R., Iwai Y., Shimoda N.,
RA Jincho Y., Hiura H., Hirose M., Mochizuki C., Yuri M., Im R-H.,
RA Funada-Wada U., Wada M.;
RT "IF3, a novel cell-differentiation factor, highly expressed in the
RT murine liver and ovary";
RL Biochem. Biophys. Res. Commun. 297:323-328(2002).
DR EMBL; AB086437; BAC11848.1; -; mRNA.
DR Ensembl; ENSMUSG00000041857; Mus musculus.
DR MGI; MGI:2149290; Oosp1.
DR GO; GO:0005615; C:extracellular space; TAS.
SQ SEQUENCE 92 AA; 10597 MW; FC458C6E10005FDA CRC64;

Alignment Scores:
Pred. No.: 0.0121 Length: 92
Score: 103.00 Matches: 29
Percent Similarity: 48.4% Conservative: 17
Best Local Similarity: 30.5% Mismatches: 31
Query Match: 12.5% Indels: 18
DB: Gaps: 6

US-10-664-025-43_COPY_53_511 (1-459) x Q8K4N9_MOUSE (1-92)
QY 7 TTAGAAGTCTTCATGCTCTCGCTGCTG-----ATTGGACCGTGCT 51
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 8 LeuGlyLeuLeuLeuLeuSerPheMetIysThrCysAlaAspTrpThrAla--- 26
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 52 GAGAACCTCAGTGAATAAGTTGCTCTCTGGACTGTTGATGCTCTCAGTTATCCCA 111
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 112 GTTGAGAAAGCAGAAATCTGTATATATTTCGGATGAATTACATCTGGGAATGGCTGC 171
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 43 ThrIle---PheHisAsnIlePheMetGluProAspGluValPheLeuGlyIleGlyCys 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 172 CCTGCAATCGGATACATACATAT-----GTATATGAGTTTATATATCTTGTTCGT 222
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 ProValThrThr-----ThrTrpProAsnAspThrTyrgluPheIleTyrgThrTyr 79
QY 223 GATTGTGCATCAGGACAGGAGTAGTTTCTGAGGAACCTCTCCTT 267
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 80 SerCysGlyIleAlaAsnLys-----SerGluSerSerPheLeu 92
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
ZP3_MESAU
ID ZP3 MESAU STANDARD; PRT; 422 AA.
AC ZP3T91.
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zona pellucida sperm-binding protein 3 precursor (Zona pellucida
DE glycoprotein ZP3) (Sperm receptor) (Zona pellucida protein C).
GN Name=ZP3; Synonyms=ZPC;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Cricetidae; Cricetinae; Mesocricetus.
```

```
OX NCBI_TaxID=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary.
RX MEDLINE=91078540; PubMed=2257975;
RA Kinloch R.A., Ruiz-Seller B., Wasserman P.M.;
RT "Genomic organization and polypeptide primary structure of zona
RT pellucida glycoprotein hZP3, the hamster sperm receptor.";
RL Dev. Biol. 142:414-421(1990).
CC -!- FUNCTION: The mammalian zona pellucida, which mediates species-
CC specific sperm binding, induction of the acrosome reaction and
CC prevents post-fertilization polyspermy, is composed of three to
CC four glycoproteins, ZP1, ZP2, ZP3, and ZP4. ZP3 is essential for
CC sperm binding and zona matrix formation.
CC -!- SUBUNIT: Polymers of ZP2 and ZP3 organized into long filaments
CC cross-linked by ZP1 homodimers (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
CC matrix upon proteolytic cleavage.
CC -!- TISSUE SPECIFICITY: Oocytes.
CC -!- DEVELOPMENTAL STAGE: Growing oocytes.
CC -!- DOMAIN: The Zp domain is involved in the polymerization of the ZP
CC proteins to form the zona pellucida.
CC -!- PTM: Proteolytically cleaved before the transmembrane segment to
CC yield the secreted ectodomain incorporated in the zona pellucida.
CC -!- PTM: N-glycosylated (By similarity).
CC -!- PTM: O-glycosylated; removal of O-linked glycans may play an
CC important role in the postfertilization block to polyspermy (By
CC similarity).
CC -!- SIMILARITY: Belongs to the ZP domain family. ZPC subfamily.
CC -!- SIMILARITY: Contains 1 ZP domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; M63629; AAA37079.1; -; mRNA.
CC InterPro; IPR001507; Endoglin/CD105.
CC Pfam; PF00100; Zona_pellucida; 1.
CC PRINTS; PR00023; ZPELUCIDA.
CC SMART; SM00241; ZP; 1.
CC PROSITE; PS00682; ZP 1; 1.
CC PROSITE; PS51034; ZP 2; 1.
CC KW Cleavage on pair of basic residues; Extracellular matrix;
KW Fertilization; Glycoprotein; Pyrrolidone carboxylic acid; Receptor;
KW Signal; Transmembrane.
FT CHAIN 1 22 By similarity.
FT PROPEP 23 349 Zona pellucida sperm-binding protein 3.
FT TOPO_DOM 23 386 Removed in mature form (By similarity).
FT TRANSMEM 387 407 Extracellular (Potential).
FT TOPO_DOM 408 422 Potential.
FT DOMAIN 45 306 Cytoplasmic (Potential).
FT COMPBAS 119 158 ZP.
FT COMPBAS 208 257 Pro-rich.
FT MOD_RES 23 23 Pyrrolidone carboxylic acid (By
FT similarity).
FT CARBOHYD 32 32 O-linked (GalNAc. .) (By similarity).
FT CARBOHYD 34 34 O-linked (GalNAc. .) (By similarity).
FT CARBOHYD 146 146 N-linked (GlcNAc. .) (By similarity).
FT CARBOHYD 155 155 O-linked (GalNAc. .) (By similarity).
FT CARBOHYD 161 161 O-linked (GalNAc. .) (By similarity).
FT CARBOHYD 162 162 O-linked (GalNAc. .) (By similarity).
FT CARBOHYD 271 271 N-linked (GlcNAc. .) (By similarity).
FT CARBOHYD 302 302 By similarity.
FT DISULFD 78 98 By similarity.
FT DISULFD 216 281 By similarity.
FT DISULFD 238 299 By similarity.
SQ SEQUENCE 422 AA; 45827 MW; D0F95BE7FF0E7E01 CRC64;

Alignment Scores:
Pred. No.: 0.294 Length: 422
```

	RA	Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B., Biemont C., Skalli Z., Cottolico L., Poulain J., De Berardinis V., Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S., Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Laudet V., Schachter V., Quettier F., Saurin W., Scarpelli C., Wincker P., Landner E.S., Weissbach J., Roest Crolius H.; "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype."; Nature 431:946-957(2004).
	RL	Nature 431:946-957(2004). [2]
	RN	[2]
	RP	NUCLEOTIDE SEQUENCE.
	RG	Genoscope; Whitehead Institute Centre for Genome Research; Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases. CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
	DR	EMBL; CAAB01011271; CAF93676.1; -; Genomic DNA. SQ SEQUENCE 408 AA; 44840 MW; CABA9142A702BD2 CRC64;
	Alignment Scores:	
Pred. No.: Length: 408		
Score: 86.50 Matches: 39		
Percent Similarity: 43.0%		
Best Local Similarity: 30.5% Mismatches: 38		
Query Match: 10.5% Indels: 35		
DB: 2 Gaps: 8		
	US-10-664-025-43_COPY_53_511 (1-459) x Q4T0A6_TETNG (1-408)	
Qy	136 ATATTGGCGATGAATACATCTGCGAAATGGCTGCCCTGCAGAACTCGGATACATCATAT 195	: : :::::
Dd	8 ValTySerAspGlueArgLeuglyVal-----GluGlnArgAspHisCysArg 24	:
Qy	196 GTAPAT-----GAGTTTATATAFTCTGT-- --CGTGAT 225	:: : :
Dd	25 AlaHisaSerSerGluGlyGlygluGluTyThrIleLeuValglyLeuAlaAasp 44	: : :::::
Qy	226 TGTGGCATCAGCACAGGGTAGTTCTCAGGAAACTCTCTTTTCAACCAGCGTGATC 285	: : :::::
Dd	45 CysgLyThryty--HisTrpValSerGluAspAlaLeuilleTyThrAsnLeulle 63	: : :::::
Qy	286 TTtAcCCCcAgG-----AAtATagATCaTGacCcTCAGGAATccAttTG 330	: : :::::
Dd	64 PheserProkrgValthrProaspglyLeulleargMetapsglualavalleProfile 83	: : :::::
Qy	331 GAGTgtTtcAcctTaGaAAA-----TCAGtgTgcCTtacCaccagtTtCTact 378	: : :::::
Dd	84 GluCysGIntyrGluArGLyetyr-SerLeuserserserserSerLeuthrProthrTrpval 103	: : :::::
Qy	379 GAGAATGNatAAAAATTTGGATCCTAGTCCTTTATTGCTGAC-----TTTCAGACA 429	: : :::::
Dd	104 -----PropheMetaAlaThrGlnAlalaValGIUthr 114	: : :::::
Qy	430 ACAGCAGAAGATTAGGATTATTA 453	: : :::::
Dd	115 LeualaPhespLeuarGleuVal 122	: : :::::
RESULT 12		
Q9C2J9 NEUCR		
ID O9C2J9 NEUCR PRELIMINARY; PRT; 1736 AA.		
AC O9C2J9;		
DT 01-JUN-2001 (TrEMBLrel. 17, Created)		
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DI 01-MAR-2004 (TrEMBRel. 26, Last annotation update)		
DE Related to regulator protein rum1.		
GN Name=3H10.130;		
OS Neurospora crassa.		
OC Eukarya; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;		
OC Sordariomycetidae; Sordariales; Sordariales; Neurospora.		
NCBI_TaxID=5141;		
RN [1]		

```

RESULT 11
Q470A6 TETNG
ID Q470A6.TETNG PRELIMINARY; PRT; 408 AA.
AC Q470A6;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Chromosome 3 SCAFl1271, whole genome shotgun sequence.
OS ORFNames=GSTENG00009437001;
GN Tetraodon nigroviridis (Green puffer) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
[1]_RN
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,

```

RESULT 13

OS Homo sapiens (Human).

RN [1]

RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,

RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,

DK EMBL; AKI27319; EAC07010.1; -, UNKNA.
DR HSSP: 001082: 1AA2.

DR GU; GU:0003779; F:actin binding; TEA.
DR Interpro: IPR001715. Calponin-like

DR SMART; SM000033; CH; 1.

Percent Similarity: 41.5% Conservative: 23

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

nb
586 r...e...f...e...g...n...t...e...a...a...d...s...n...v...s...t...e...g...n...a...s...n...g...l...y...a...t...e...u...a...s...n...c...y...s...g...l...u...g...l...u... 605

QY 67 AAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCCAGTTGCAGAAAGCAGA 126

OV 127 AATCTGTATATATTGCGGATGAATTACATCTGGGAATGGCTGCCCTGCA----- 177

Db 616 -----AlaAspAlaAlaHisLeuGluSerGlyGlnProValGlnCysGlu 630

QY 178 AACGGATACATACATATGTAATGAGTITATATATCTTGTCGTGATCTGCGCATCAGG 23

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```

Db      531  HisPhThrSerLeuGluLeuValProProSerThrLeuThrThrHisLeuLysAla 550
Qy      397  GATCCT 402
          ::|||
Db      551  GluPro 552

Search completed: May 9, 2006, 10:09:19
Job time : 122.9 secs

```

Search completed: May 9, 2006, 10:09:19
Job time : 122.9 secs

Alignment Scores:	
Pred. No.:	3.65
Score:	85.50
Percent Similarity:	41.5%
Best Local Similarity:	25.4%
Query Match:	10.4%
DB:	2
Gaps:	5
Indels:	25
Mismatches:	58
Conservative:	23
Matches:	36
Length:	5430

US-10-664-025-43 COPY 53 511 (1-459) x Q5VW20 HUMAN (1-5430)

[illegible]

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search using frame plug_p2n model

Run on: May 9, 2006, 13:15:42 ; Search time 4123.8 Seconds
(without alignments)
3163.483 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLAVLIWTGAENL.....PSPFIADFTAAELGLLIF 153

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 588141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/abs/ABSWEB/spool/US10664025/runat_08052006_173641_29351/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p
-USERS=US10664025 @CGN_1_1724 @runat_08052006_173641_29351 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sta:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	799	100.0	512	6	BD107959 EST and e
2	799	100.0	512	6	AR412406 Sequence
3	799	100.0	512	6	AX969240 Sequence

4	789	98.7	477	6	CQ737424	CQ737424 Sequence
5	789	98.7	642	6	BD109278	BD109278 EST and e
6	789	98.7	642	6	AR413725	AR413725 Sequence
7	789	98.7	642	6	AX970559	AX970559 Sequence
8	789	98.7	691	6	CS072271	CS072271 Sequence
9	789	98.7	1337	6	AX574405	AX574405 Sequence
10	789	98.7	1480	6	AX600204	AX600204 Sequence
11	789	98.7	1502	8	BC036256	BC036256 Homo sapi
12	789	98.7	1646	8	BC048121	BC048121 Homo sapi
13	789	98.7	1898	6	BD275948	BD275948 62 Human
14	785	98.2	1492	6	AX748179	AX748179 Sequence
15	785	98.2	1492	6	AK093517	AK093517 Homo sapi
16	775	97.0	497	6	AX588689	AX588689 Sequence
17	752	94.1	500	6	AX588688	AX588688 Sequence
18	608	76.1	560	6	BD110503	BD110503 EST and e
19	608	76.1	560	6	AR414950	AR414950 Sequence
20	608	76.1	560	6	AX971784	AX971784 Sequence
21	468	58.6	372	6	BD119756	BD119756 EST and e
22	468	58.6	372	6	AR424203	AR424203 Sequence
23	468	58.6	372	6	AX984897	AX984897 Sequence
24	431.5	54.0	655	9	BC099498	BC099498 Mus muscu
25	429	53.7	469	6	BD109279	BD109279 EST and e
26	429	53.7	469	6	AR413726	AR413726 Sequence
27	429	53.7	469	6	AX970560	AX970560 Sequence
28	429	53.7	470	6	CS072276	CS072276 Sequence
29	325.5	40.7	100317	14	AF000409	AF000409 Homo sapi
30	325.5	40.7	101031	14	AF000622	AF000622 Homo sapi
31	325.5	40.7	176944	8	AP000790	AP000790 Homo sapi
32	325.5	40.7	18367	8	AC013807	AC013807 Homo sapi
33	325	40.7	10115	6	AX574404	AX574404 Sequence
34	240	30.0	139212	14	AC154077	AC154077 Loxodonta
35	234	29.3	1074	9	AF234653	AF234653 Mus muscu
36	234	29.3	1096	9	AF250838	AF250838 Mus muscu
37	234	29.3	110000	14	AL845451	Continuation (3 of
38	234	29.3	195039	9	BX649621	BX649621 Mouse DNA
39	230	28.8	1075	9	BC051666	BC051666 Mus muscu
40	224.5	28.1	589	6	AX361328	AX361328 Sequence
41	223.5	28.0	1067	8	BC066327	BC066327 Homo sapi
42	223.5	28.0	1126	6	BD127476	BD127476 Primer fo
43	223.5	28.0	1126	6	CQ783102	CQ783102 Sequence
44	223.5	28.0	1126	8	AK075086	AK075086 Homo sapi
45	223.5	28.0	1131	6	CS050200	CS050200 Sequence

ALIGNMENTS

RESULT 1	BD107959	BD107959	512 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD107959	EST and encoded human protein.				
DEFINITION	BD107959	EST and encoded human protein				
ACCESSION	BD107959	BD107959				
VERSION	BD107959.1	GI:23202777				
KEYWORDS	JP 2002010789-A/36					
SOURCE	Homo sapiens	Homo sapiens (human)				
ORGANISM	Homo sapiens	Homo sapiens				
REFERENCE	1	(bases 1 to 512)				
AUTHORS	Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.					
TITLE	EST and encoded human protein					
JOURNAL	Patent: JP 2002010789-A 36 15-JAN-2002;					
COMMENT	GENSET CORP					
	OS Homo sapiens (human)					
	PN JP 2002010789-A/36					
	PD 15-JAN-2002					
	PF 07-AUG-2000 JP 2002080989					
	PR 05-AUG-1999 US 60/147499					
	PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE					
	GIORDANO					
	PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC					
	C12N1/21,					
	PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC					

```
C12N15/00
CC Von Heijne matrix
CC score 10.6999998092651
CC seq VMLLAVLWGA/EN
FH key Location/Qualifiers
FT CDS 53..511
FT sig_peptide 53..103.
FEATURES
    source
        1..512
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
    ORIGIN
        US-10-664-025-3903 (1-153) x BD107959 (1-512)
        Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrrThrGlyAlaGluAsnLeu 20
        Db 53 ATGGCGTTAGAAATCTCTCGATGCTCTCGCTGCTTGGATTGGACCGGTGCTGAGAACCTC 112
        Qy 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
        Db 113 CATGTGAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 172
        Qy 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
        Db 173 AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGCTGCCCTCAAAAT 232
        Qy 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
        Db 233 CGGATACATACATATGATATATAGTTATATATCTTTGTTGGATGTGGCATCAGGACA 292
        Qy 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
        Db 293 AGGTAGTTTCTGAGAAACTCTCTTTTCAACCGAGCTGACTTTACCCCAAGGAAT 352
        Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
        Db 353 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 412
        Qy 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
        Db 413 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAATTTGGATCTTAGTCTTTATTGCTGAC 472
        Qy 141 PheGlnThrThrAlaGluLeuGlyLeuLeuIlePhe 153
        Db 473 TTTCAGACAACAGCAGAGTAGGATTATTAATCTTC 511
    RESULT 2
    LOCUS AR412406 512 bp DNA linear PAT 18-DEC-2003
    DEFINITION Sequence 43 from patent US 6639063.
    ACCESSION AR412406
    VERSION AR412406.1 GI:40167516
    KEYWORDS
    SOURCE
    ORGANISM
        Unknown.
        Unclassified.
    REFERENCE
        1 (bases 1 to 512)
        Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
        EST's and encoded human proteins
    TITLE
        Patent: US 6639063-A 43 28-OCT-2003;
    JOURNAL
        Genet. S.A.;
    FEATURES
        WOX; Location/Qualifiers
```

```
source
    1..512
    /organism="unknown"
    /mol_type="genomic DNA"
    ORIGIN
        Alignment Scores:
        Pred. No.: 3.49e-85 Length: 512
        Score: 799.00 Matches: 153
        Percent Similarity: 100.0% Conservative: 0
        Best Local Similarity: 100.0% Mismatches: 0
        Query Match: 100.0% Indels: 0
        DB: 6 Gaps: 0
        US-10-664-025-3903 (1-153) x AR412406 (1-512)
        Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrrThrGlyAlaGluAsnLeu 20
        Db 53 ATGGCGTTAGAAATCTCTCGATGCTCTCGCTGCTTGGATTGGACCGGTGCTGAGAACCTC 112
        Qy 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
        Db 113 CATGTGAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 172
        Qy 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
        Db 173 AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGCTGCCCTCAAAAT 232
        Qy 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
        Db 233 CGGATACATACATATGATATATAGTTATATATCTTTGTTGGATGTGGCATCAGGACA 292
        Qy 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
        Db 293 AGGTAGTTTCTGAGAAACTCTCTTTTCAACCGAGCTGACTTTACCCCAAGGAAT 352
        Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
        Db 353 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 412
        Qy 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
        Db 413 CTTACACCAAGTTTCTACTGAGAAATGAAATAAAATTTGGATCTTAGTCTTTATTGCTGAC 472
        Qy 141 PheGlnThrThrAlaGluLeuGlyLeuLeuIlePhe 153
        Db 473 TTTCAGACAACAGCAGAGTAGGATTATTAATCTTC 511
    RESULT 3
    LOCUS AX969240 512 bp DNA linear PAT 15-JAN-2004
    DEFINITION Sequence 43 from Patent EP1104808.
    ACCESSION AX969240
    VERSION AX969240.1 GI:40975399
    KEYWORDS
    SOURCE
    ORGANISM
        Homo sapiens (human)
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Homindae; Homo.
    REFERENCE
        1
        Dumas Milne Edwards, J.B., Jobert, S. and Giordano, J.Y.
        ESTs and encoded human proteins
    TITLE
        Patent: EP 1104808-A 43 06-JUN-2001;
    JOURNAL
        Genset (FR)
    FEATURES
        Location/Qualifiers
        1..512
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        53..>511
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAF14348.1"
```

/db_xref="GI:40975400"
/translations="MALEVLMLAVLIVITGAENLHVKISCSLDMLVSVTPVAESNRL
YFADLEHLGCMGPANRIHYVEFLYLPVDCIRXRVVSEETLLFQTELYFTFPRNID
HDPQEIHLCESTRKSWLTPVSTENIKLDPSPFFIADFQTTAAEEIUGLLIF"
53...103
/notes="Von Heijne matrix score 10.69999998092651 seq
VLMLLAVLITGA/EN"

sig_peptide

ORIGIN

Alignment Scores:

Pred. No.: 3,49e-85 Length: 512
Score: 799.00 Matches: 153
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x AX969240 (1-512)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleThrThrGlyAlaGluAAsnLeu 20
DB 53 ATGGCGTTAGAGTCTTGATCTCTCGCTGCTTGTGATTTGGACCGGTGAGAACCTC 112
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 113 CATGTCAAAATAAGTTGCTCTCTGGACTGTTGATGCTCAGTTATCCAGTTGAGAA 172
QY 41 SerArgAsnLeuTyrrilePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGATGGCTGCCCTGCAAT 232
QY 61 ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 292
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
DB 293 AGGGTAGTTCTTGAGGAAATCTCTCTTTTCAAAACCGAGCTGTACTTTACCCCAAGG 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
DB 353 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACCAAGTTCTTACTCAGAAATGAATAAAATTGGATCCTAGTCTTTTATTGCTGAC 472
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeuPhe 153
DB 473 TTTGAGCAACAGCAGAGAGTTAGGATTTATTAATCTTC 511

RESULT 4

LOCUS CQ737424
DEFINITION Sequence 23358 from Patent WO02068579.
ACCESSION CQ737424
VERSION CQ737424.1 GI:42335713
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 23358 06-SEP-2002;
PE Corporation (NY) (US)
LOCATION/Qualifiers
1. 477
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 4.92e-84 Length: 477
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x CQ737424 (1-477)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleThrThrGlyAlaGluAAsnLeu 20
DB 1 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGAGAACCTC 60
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 61 CATGTCAAAATAAGTTGCTCTCTGGACTGTTGATGCTCAGTTATCCAGTTGAGAA 120
QY 41 SerArgAsnLeuTyrrilePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 121 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGATGGCTGCCCTGCAAT 180
QY 61 ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr 80
DB 181 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
DB 241 AGGGTAGTTCTTGAGGAAATCTCTCTTTTCAAAACCGAGCTGTACTTTACCCCAAGG 300
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
DB 301 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 360
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 361 CTTACACCAAGTTCTTACTCAGAAATGAATAAAATTGGATCCTAGTCTTTTATTGCTGAC 420
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
DB 421 TTTGAGCAACAGCAGAGAGTTAGGATTTATTA 453

RESULT 5

LOCUS BD109278
DEFINITION EST and encoded human protein.
ACCESSION BD109278
VERSION BD109278.1 GI:23204096
KEYWORDS JP 2002010789-A/1355.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 1355 15-JAN-2002;
GENSET CORP

COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/1355
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE
GIORDANO

PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,PC
C12N1/21,
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00,PC

```

C12N15/00
CC Von Heijne matrix
CC score 10.699998092651
CC seq VMLLAVALIWTGA/EN
FH key Location/Qualifiers
FT CDS 53..526
FT sig_peptide 53..103.
FEATURES
    source
        1..642
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7,048-84 Length: 642
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0
US-10-664-025-3903 (1-153) x BD109278 (1-642)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
DB 53 ATGGCGTTAGAACTCTCGATGCTCTCGCTGCTTGTGATTGGACCGGTGCTGAGAACCTC 112
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 113 CATGTGAAAATAAGTTGCTCTCTGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 172
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAAATCTGTATATATTTGCGATGAATTACATCTCGGAATGGGCTGCCCTGCAAT 232
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATGATGATGATATATATCTTGTTCGGATGTGGCATCAGGACA 292
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 293 AGGGTAGTTTCTGAGAAACTCTCCTTTTCAACCGAGCTGCTACTTTACCCCAAGGAAT 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleSerValTrp 120
DB 353 ATAGATCATGACCCCTCAGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCTTAGTCTCTTTATTGCTGAC 472
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
DB 473 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 505
RESULT 6
LOCUS AR413725 642 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1362 from patent US 6639063.
ACCESSION AR413725
VERSION AR413725.1 GI:40168835
KEYWORDS
SOURCE
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE
    1
    Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.-Y.
    ESTs and encoded human proteins
    Patent: EP 1362 28-OCT-2003;
    Genset S.A.;
    WOX;
    Location/Qualifiers
        1..642
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        53..>526
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAF15667.1"
FEATURES
    source
        1..642
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        53..>526
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAF15667.1"

```

```

1..642
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 7,048-84 Length: 642
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0
US-10-664-025-3903 (1-153) x AR413725 (1-642)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
DB 53 ATGGCGTTAGAACTCTCGATGCTCTCGCTGCTTGTGATTGGACCGGTGCTGAGAACCTC 112
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 113 CATGTGAAAATAAGTTGCTCTCTGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 172
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAAATCTGTATATATTTGCGATGAATTACATCTCGGAATGGGCTGCCCTGCAAT 232
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATGATGATGATATATCTTGTTCGGATGTGGCATCAGGACA 292
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 293 AGGGTAGTTTCTGAGAAACTCTCCTTTTCAACCGAGCTGCTACTTTACCCCAAGGAAT 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleSerValTrp 120
DB 353 ATAGATCATGACCCCTCAGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACAGTTTCTACTGAGATGAATAAATTTGGATCTTAGTCTCTTTATTGCTGAC 472
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
DB 473 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 505
RESULT 7
LOCUS AX970559 642 bp DNA linear PAT 15-JAN-2004
DEFINITION Sequence 1362 from Patent EP1104808.
ACCESSION AX970559
VERSION AX970559.1 GI:40978037
KEYWORDS
SOURCE
ORGANISM
    Homo sapiens (human)
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE
    1
    Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.-Y.
    ESTs and encoded human proteins
    Patent: EP 1104808-A 1362 06-JUN-2001;
    Genset (FR)
    Location/Qualifiers
        1..642
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        53..>526
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAF15667.1"
CDS

```


/db xref="GI:40978038"
/translation="MALEVLMLAVLITWTGAENLHVKISCSLDLWLVSVIPVAESRNL
YIYADELHLMGCPANRIHYTFYFYLVRDGGIRTRVSEETLFTQTELYFTFRNID
HDPQELHLCSTSRKSVMLTPVSTENEIKLDPSPFIADPFTTAAEELGLSSPNLL"
53. .103
/note="Von Heijne matrix score 10.6999998092651 seq
VLMLLAVLITWGA/EN"

sig_peptide

ORIGIN

Alignment Scores:
Pred. No.: 7.04e-84 Length: 642
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x AX970559 (1-642)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAAsnLeu 20
DB 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTTGGACCGGTGCTGAGAACCTC 112
QY 21 HisValIysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 113 CATGTGAAATAAGTTGCTCTCGACTGGTGGTGGTCTCAGTTATCCAGTTGCGAGAA 172
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGCTGCCCTGCAAA 232
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATGTATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 292
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 293 AGGGTAGTTCTTGAGGAACTCTCTTTTCAACCGAGCTGACCTTACCCCAAGAA 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTyr 120
DB 353 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTG 412
QY 121 LeuThrProValSerThrGluAnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACCAAGTTCTACTGAGATGAATAAAATGGATCCTAGTCTTATTTATGCTGAC 472
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
DB 473 TTTCAGACACAGCAGAGTTAGGATTATTA 505

RESULT 8
CS072271
LOCUS
DEFINITION
Sequence 19 from Patent WO2001042451.
ACCESSION
CS072271
VERSION
CS072271.1 GI:63089501
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1
AUTHORS
Dumas Milne Edwards,J.B., Bougueleret,L. and Jobert,S.
TITLE
Full-length human cDNAs encoding potentially secreted proteins
JOURNAL
Patent: WO 2001042451-A 19 14-JUN-2001;
Seron Genetics Institute S.A. (FR)

FEATURES
1. .691
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS

42. .518
/note="unnamed protein product"
/codon start=1
/protein_id="CAI93409.1"
/db xref="GI:63089502"
/translation="MALEVLMLAVLITWTGAENLHVKISCSLDLWLVSVIPVAESRNL
YIYADELHLMGCPANRIHYTFYFYLVRDGGIRTRVSEETLFTQTELYFTFRNID
HDPQELHLCSTSRKSVMLTPVSTENEIKLDPSPFIADPFTTAAEELGLSSPNLL"
42. .92
/note="Von Heijne matrix score 10.7019149919754 seq
VLMLLAVLITWGA/EN"

sig_peptide

ORIGIN

Alignment Scores:
Pred. No.: 7.69e-84 Length: 691
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x CS072271 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAAsnLeu 20
DB 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGGATTTGGACCGGTGCTGAGAACCTC 101
QY 21 HisValIysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 102 CATGTGAAATAAGTTGCTCTCGACTGGTGGTGGTCTCAGTTATCCAGTTGCGAGAA 161
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 162 AGCAGAAATCTGTATATATTGGGATGAATTACATCTGGGAATGGCTGCCCTGCAAA 221
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 222 CGGATACATACATATGTATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 281
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 282 AGGGTAGTTCTTGAGGAACTCTCTTTTCAACCGAGCTGACCTTACCCCAAGAA 341
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTyr 120
DB 342 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTG 401
QY 121 LeuThrProValSerThrGluAnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 402 CTTACACCAAGTTCTACTGAGATGAATAAAATGGATCCTAGTCTTATTTATGCTGAC 461
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
DB 462 TTTCAGACACAGCAGAGTTAGGATTATTA 494

RESULT 9
AX574405
LOCUS
DEFINITION
Sequence 2 from Patent WO02060466.
ACCESSION
AX574405
VERSION
AX574405.1 GI:27551738
KEYWORDS
SOURCE
Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1
AUTHORS
Salter-Cid,L., Ebbets-Reed,D., Bour,B.A., Chicca,J., Yen-Potin,F.
and Bihain,B.
TITLE
Gasp3 polynucleotides and polypeptides and uses thereof
JOURNAL
Patent: WO 02060466-A 2 08-AUG-2002;
GENSET (FR)

FEATURES	Location/Qualifiers		Proteins associated with cell growth, differentiation, and death	
	1..1337		Patent: WO 02072830-A 16 19-SEP-2002;	
	/organism="Homo sapiens"		Incyte Genomics, Inc. (US)	
	/mol_type="unassigned DNA"		Location/Qualifiers	
CDS	/db_xref="taxon:9606"		1..1480	
	53..>103		/organism="Homo sapiens"	
	/note="unnamed protein product"		/mol_type="unassigned DNA"	
	/codon_start=1		/db_xref="taxon:9606"	
sig_peptide	/protein_id="CAD60066.1"		/note="Incyte ID No: 5284076CB1"	
	/db_xref="GI:27551739"			
	/translation="MADEVLMMLAVLIWTGA"			
	104..526			
ORIGIN				
Alignment Scores:				
Pred. No.:	1.7e-83	Length:	1337	1480
Score:	789.00	Matches:	151	151
Percent Similarity:	100.0%	Conservative:	0	0
Best Local Similarity:	100.0%	Mismatches:	0	0
Query Match:	98.7%	Indels:	0	0
DB:	6	Gaps:	0	0
US-10-664-025-3903 (1-153) x AX574405 (1-1337)				
Qy	1	MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTTrpThrGlyAlaGluAsnLeu	20	
Db	53	ATGGCGTTAGAACTTTGATGCTCTCGCTGCTTTGATTTGGACCGGTGCTGAGAACCTC	112	
Qy	21	HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40	
Db	113	CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA	172	
Qy	41	SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn	60	
Db	173	AGCAGAAATCTGTATATATTTCCGATGAATACATCTGGGAATGGCGTGCCTCGCAAT	232	
Qy	61	ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr	80	
Db	233	CGGATACATACATATGATATAGTTTATATATCTTGTTCGATTTGGCATCAGGACA	292	
Qy	81	ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn	100	
Db	293	AGGCTAGTTTCTGAGGAAACTCTCTTTTCAACCCGAGCTGTACTTTACCCCAAGGAAT	352	
Qy	101	IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp	120	
Db	353	ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG	412	
Qy	121	LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp	140	
Db	413	CTTACACCACTTCTACTGAGAAATGAATAAATTTGGATCCTAGTCTCTTTATTGCTGAC	472	
Qy	141	PheGlnThrThrAlaGluGluLeuGlyLeuLeu	151	
Db	473	TTTCAGACAAACGACAGAGAGTTAGGATTATTA	505	
RESULT 10				
AX600204				
LOCUS				
DEFINITION	Sequence 16 from Patent WO02072830.			
ACCESSION	AX600204			
VERSION	AX600204.1 GI:28400246			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
AUTHORS	Hominidae; Homo.			
1	(bases 1 to 1502)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,				
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.				
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,				
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,				
Diatchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L.,				
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,				

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Bosak, S.A., McEwan, P.J., Abramson, R.D., Mullahy, S.J., Moore, T., Max, S.I., Wang, J., Hsieh, P., McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Worley, K.C., Halse, S., Sodergren, E.J., Lu, X., Gibbs, R.A., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1502)

NIH MGC Project
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 74 Row: K Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 42476002.

FEATURES

Source

1..1502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:43628 IMAGE:5270384"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescriptR"
1..1502
/gene="FLJ36198"
/db_xref="GeneID:219990"
52..528
/gene="FLJ36198"
/codon_start=1
/product="hypothetical protein LOC219990"
/protein_id="AAH36256.1"
/db_xref="GI:37537253"
/db_xref="GeneID:219990"
/translation="MALEVLMLGLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNL YIFADLHLMGCPANRHTYEFYILVRDCGIRTVVSEETLFPQELYPTRNID HDPOIHLCESTRKSVMLTPVSTENEIKLDPSPTIADPQTAAELGLLSGPNLL"

ORIGIN

Alignment Scores:
Pred. No.: 1,96e-83 Length: 1502
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0

DB: 8 Gaps: 0
US-10-664-025-3903 (1-153) x BC036256 (1-1502)
QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuIleTrpThrGlyAlaGluAAsnLeu 20
52 ATGGCGTTAGAACTCTTGATGCTCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 111
QY 21 HisValIysIleSerCysSerLeuAAspTrpLeuMetValSerValIleProValAlaGlu 40
112 CATGTGAAAATAGATTGCTCTCTGGACTGCTGTTGATGCTCTCAGTTATCCCAAGTTCAGAA 171
QY 41 SerArgAsnLeuTrpIlePheAlaAAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
172 AGCAGAAATCTGATATATATTTGGGATGAATACATCTGGGAATGGGCTGCCCTGCAAA 231
QY 61 ArgIleHisThrTrpValTrpGluPheIleTrpLeuValArgAAspCysGlyIleArgThr 80
232 CGGATACATACATATGAT 291
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTrpPheThrProArgAsn 100
292 AGGGTAGTTTCTGAGGAAACTCTCTCTTTTCAAAACGAGCTGTACTTTACCCCAAGGAAT 351
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTrp 120
352 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACACCTCTAGGAAATCAGTGTGG 411
QY 121 LeuThrProValSerThrGluAAsnGluIleLysLeuAAspProSerProPheIleAlaAAsp 140
412 CTTACACCAAGTTTCTACTGAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 471
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
472 TTTTCAGACACACGACGAGAGTAGGATTATTA 504

RESULT 12
BC048121 1646 bp mRNA linear PRI 28-JUL-2005
Homo sapiens hypothetical protein FLJ36198, mRNA (cdna clone
MGC:57403 IMAGE:5266642), complete cds.
BC048121
ACCESSION BC048121.1 GI:28856186
VERSION
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1646)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, A.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Bosak, S.A., McEwan, P.J., Abramson, R.D., Mullahy, S.J., Sodergren, E.J., Lu, X., Gibbs, R.A., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

```

REFERENCE 2 (bases 1 to 1646)
AUTHORS NIH MGC Project
CONSTRM Direct Submission
TITLE Submitted (04-MAR-2003) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 106 Row: j Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 42476002.
FEATURES
source
    1..1646
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="MGC:57403 IMAGE:5266642"
        /tissue_type="Testis"
        /clone_lib="NIH MGC_97"
        /lab_host="DH10B"
        /note="Vector: pBluescriptR"
gene
    1..1646
        /gene="FLJ36198"
        /db_xref="GeneID:219990"
        186..662
        /gene="FLJ36198"
        /codon_start=1
        /product="hypothetical protein LOC219990"
        /protein_id="AAH48121.1"
        /db_xref="GI:28856187"
        /db_xref="GeneID:219990"
        /translation="MALSLVLMIAVLITWGAENLHVKISCLDMLMVSVIPVAESRNL
        YIFADELHMGCPANRITHVYEFILVRDCGTRVVSSETLLFQELVETPRNID
        HDPQEIHLCESTRKSVWLTVPVSTENIKLDPSPFIADQITTAELGLLSSPNLL"
ORIGIN
Alignment Scores:
Pred. No.: 2,198-83 Length: 1646
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservat: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x BC048121 (1-1646)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrrThrGlyAlaGluAenLeu 20
Db 186 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 245
QY 21 HisValIysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValaGlu 40
Db 246 CATGTGAAATAAAGTTGCTCTCTGACTGGTTGTGATGCTCAGTTATCCCAATTCAGAA 305
QY 41 SerArgAsnLeuTyIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 306 AGCAGAAATCTGTATATATTTCGGATGNAATTACATCTGGGAATGGCGTGCCTGCAAT 365
QY 61 ArgIleHisThrTyValTrpGluPheIleTyLeuValArgAspCysGlyIleArgThr 80

```

```

Db 366 CGGATACATACATATGATATATGAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 425
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyPheThrProArgAsn 100
Db 426 AGGGTAGTTTCTGAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 485
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIysSerValTrp 120
Db 486 ATAGATCATGACCCCTCAGGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 545
QY 121 LeuThrProValSerThrGluAsnGluIleIysLeuAspProSerProPheIleAlaAsp 140
Db 546 CTTACACAGTTTCTTACTGAGAAATGAATAAAATTTGGATCTCTAGTCCTTTTATTCTCTGAC 605
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db 606 TTTGACACACAGCAGAGAGTTAGGATTATTA 638

RESULT 13
BD275948 1898 bp DNA linear PAT 17-JUL-2003
LOCUS 62 Human secreted proteins.
DEFINITION BD275948
ACCESSION BD275948.1 GI:33085716
VERSION JP 2002543771-A/26.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1898)
AUTHORS Birse, C.E., Mouret, P.A., Florence, K.A., Ruben, S.M.,
Komatsoulis, G.A., Ni, J., Ebner, R., W.D., Lafleur, Olsen, H.S.,
Shi, Y., Soppet, D.R., Rosen, C.A. and Young, P.E.
62 Human secreted proteins
Patent: JP 2002543771-A 26 24-DEC-2002;
Human Genome Sciences Inc
COMMENT OS Homo sapiens
PN JP 2002543771-A/26
PD 24-DEC-2002
PF 06-APR-2000 JP 2000611564
PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
Charles, e birse,paul a mouret,kimberly a florence,steven m pi
ruben,
PI george a komatsoulis,jian ni,reinhard ebner,david w pi
lafleur,henrik e olsen
PI yanggu shi,daniel r soppet,craig a rosen,paul e young CC
FH Key Location/Qualifiers
FT SITE (1398)
FT SITE (1428)
FT SITE /note='n equals a,t,g, or c'
FT SITE /note='n equals a,t,g, or c'
FEATURES
source
    1..1898
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2,6e-83 Length: 1898
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservat: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x BD275948 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrrThrGlyAlaGluAenLeu 20
Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 74

```

QY 21 HisVallylsleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
 DB 75 CATGTGAANAATAAGTTGCTCTCTGGACTGGTTGATGCTCAGTTATCCCGAGTTGCAGAA 134
 QY 41 SerArgenLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
 DB 135 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTGGGAATGGCTGCCCTGCCAAAT 194
 QY 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80
 DB 195 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA 254
 QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
 DB 255 AGGTAGTTCTTCAGGAAATCTCTCTTTTCAAAACCGAGCTGTACTTTACCCCAAGGAAT 314
 QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArglyssSerValTrp 120
 DB 315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 374
 QY 121 LeuThrProValSerThrGluAsnGluIleHisLeuGluCysSerThrSerArglyssSerValTrp 140
 DB 375 CTTACACCACTTCTACTGAGAATGAATAAATTGGATCCTAGTCTCTTTATTGCTGAC 434
 QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
 DB 435 TTTACAGACAACAGAGAAGTAGTTAGGATTATTA 467
 RESULT 14
 LOCUS AX748179 1492 bp mRNA linear PAT 20-JUN-2003
 DEFINITION Sequence 1704 from Patent EP1308459.
 ACCESSION AX748179
 VERSION AX748179.1 GI:32132567
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.
 TITLE Full-length cDNA sequences
 JOURNAL Patent: EP 1308459-A 1704 07-MAY-2003;
 Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
 FEATURES
 source Location/Qualifiers
 1..1492
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.8e-83 Length: 1492
 Score: 785.00 Matches: 150
 Percent Similarity: 99.3% Conservative: 0
 Best Local Similarity: 99.3% Mismatches: 1
 Query Match: 98.2% Indels: 0
 DB: 6 Gaps: 0
 US-10-664-025-3903 (1-153) x AX748179 (1-1492)
 QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
 DB 52 ATGGCGTTAGAAGCTTGTGATGCTCTCGCTGCTTGTGATTGGACCGGTCTGAGAACCTC 111
 QY 21 HisVallylsleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
 DB 112 CATGTGAANAATAAGTTGCTCTCTGGACTGGTTGATGGCTCAGTTATCCCGAGTTGCAGAA 171

QY 41 SerArgenLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
 DB 172 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTGGGAATGGCTGCCCTGCCAAAT 231
 QY 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80
 DB 232 CGGATACATACATATGTATATAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA 291
 QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
 DB 292 AGGTAGTTCTTCAGGAAATCTCTCTTTTCAAAACCGAGCTGTACTTTACCCCAAGGAAT 351
 QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArglyssSerValTrp 120
 DB 352 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 411
 QY 121 LeuThrProValSerThrGluAsnGluIleHisLeuGluCysSerProSerProPheIleAlaAsp 140
 DB 412 CTTACACCACTTCTACTGAGAATGAATAAATTGGATCCTAGTCTCTTTATTGCTGAC 471
 QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
 DB 472 TTTACAGACAACAGAGAAGTAGTTAGGATTATTA 504
 RESULT 15
 LOCUS AK093517 1492 bp mRNA linear PRI 30-JAN-2004
 DEFINITION to Mus musculus BPC526 mRNA.
 ACCESSION AK093517
 VERSION AK093517.1 GI:21752414
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oyayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
 TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 JOURNAL 14702039
 PUBMED 2
 REFERENCE

AUTHORS

Ishibaishi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
Ono, Y., Hotta, T., Hiraoka, S., Murakawa, K., Takiguchi, S.,
Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y.,
Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE

NEDO human cDNA sequencing project

JOURNAL

Unpublished

REFERENCE

3 (bases 1 to 1492)

AUTHORS

Isogai, T. and Yamamoto, J.

TITLE

Direct Submission

JOURNAL

Submitted (04-JUL-2002)

Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

Source

1..1492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI2028242"
/tissue_type="testis"
/clone_lib="TESTI2"
/note="cloning vector: pME18SFL3"
52..528
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04191.1"
/db_xref="GI:21752415"
/translation="MALEVLMLLAVLIWTGAENLHVKJSCSLDWLMAVLPVASESNL
YIFABELHLMGCPANRIHTYVFIYLRDCGIRTVVSEETLFFQELVFTPRNID
HDPQBIHLCESTSRKSVMLTPVSTENEIKLDPSPFIADPQTITAEELGLLSSSPNLL"

CDS

Alignment Scores:

Pred. No.:	5,8e-83	Length:	1492
Score:	785.00	Matches:	150
Percent Similarity:	99.3%	Conservative:	0
Best Local Similarity:	99.3%	Mismatches:	1
Query Match:	98.2%	Indels:	0
DB:	8	Gaps:	0

US-10-664-025-3903 (1-153) x AK093517 (1-1492)

QY	1	MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu	20
Db	52	ATGGCGTTAGAGTCTTGATGCTCCTCGCTGCTTGATTTGGACCGGTCGAGAACCTC	111
QY	21	HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
Db	112	CATGTGAAATAAAGTTGCTCTGGACTGGTTGATGGCCCTCAGTTATCCAGTTGCAGAA	171
QY	41	SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn	60
Db	172	ACGAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT	231
QY	61	ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr	80
Db	232	CGGATACATACATATGATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA	291
QY	81	ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn	100
Db	292	AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAACCGAGCTGCTACTTTTACCCCAAGGAT	351

QY	101	IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp	120
Db	352	ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG	411
QY	121	LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp	140
Db	412	CTTACACCAAGTTTCTACTGAGAAATGAAATGGAATTCCTAGTCTCTTTTATTGCTGAC	471
QY	141	PheGlnThrThrAlaGluGluLeuGlyLeuLeu	151
Db	472	TTTCAGACAACAGCAGAAAGATTAGGATTATTATA	504

Search completed: May 9, 2006, 14:49:44
Job time : 4129.8 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 13:15:10 ; Search time 795.6 Seconds

(without alignments)

1922.507 Million cell updates/sec

Title:

US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLWTGAENL.....PSPFIADPQTABELGLLIF 153

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delpop 6.0 , Delcxt 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/abs/ABSSWEB_spool/US10664025/runat_08052006_173639_29334/app_query.fasta 1

-DB=N Geneseq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs07

-USER=US10664025 -CGEN_1_1_900 @runat_08052006_173639_29334 -NCPU=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	789	98.7	691	5 AAH64743	Aah64743 Human sec
2	789	98.7	1337	6 ABK90053	Abk90053 DNA encod
3	789	98.7	1480	6 ABS78646	Abs78646 Human cDN
4	789	98.7	1898	3 AAC69528	Aac69528 Human sec

5	789	98.7	1898	8 ACC50642	Acc50642 Human sec
6	789	98.7	1898	8 ABZ71354	Abz71354 Secreted
7	789	98.7	1898	9 ADB91291	Adb91291 Human sec
8	789	98.7	1898	10 ADC73723	Adc73723 Human sec
9	785	98.2	1492	10 ADB63550	Adb63550 Human cDN
10	775	97.0	497	8 ABZ36691	Abz36691 Human GEN
11	752	94.1	500	8 ABZ36690	Abz36690 Human GEN
12	429	53.7	470	5 AAH64748	Aah64748 Human sec
13	325.5	40.7	5013	12 ADJ12678	Adj12678 DNA fragm
14	325	40.7	10115	6 ABK90052	Abk90052 Human GSS
15	225.5	28.2	838	10 ADE28662	Ades28662 Human NOV
16	225.5	28.2	838	12 ADM93407	Adm93407 Human NOV
17	224.5	28.1	589	6 ABL92990	AbL92990 Rat metas
18	224.5	28.1	589	12 ADN07781	Adn07781 Human mam
19	223.5	28.0	1126	4 AAK94447	Aak94447 Human ful
20	223.5	28.0	1126	12 ADL31209	Adl31209 Full leng
21	223.5	28.0	1131	14 ADY52669	Ady52669 Human fet
22	223.5	28.0	1148	2 AAX51726	Aax51726 DNA encod
23	223.5	28.0	1148	6 ABQ92578	Abq92578 Human sec
24	223.5	28.0	1153	2 AAV55748	Aav55748 Human sec
25	223.5	28.0	1153	6 ABQ92074	Abq92074 Human pol
26	223.5	28.0	1187	3 AAA26978	Aaa26978 Human cor
27	223.5	28.0	1187	10 ADC24643	Adc24643 Human cDN
28	204.5	25.6	784	4 AAK92136	Aak92136 Human cDN
29	204.5	25.6	784	4 AAK93954	Aak93954 Human cDN
30	204.5	25.6	784	12 ADL30381	Adl30381 5' end of
31	204.5	25.6	784	12 ADL28563	Adl28563 5' end of
32	201.5	25.2	646	4 AAK93399	Aak93399 Human cDN
33	201.5	25.2	646	12 ADL29826	Adl29826 5' end of
34	183.5	23.0	420	10 ACF57862	AcF57862 PLAC1 int
35	167	20.9	386	2 AAV87718	Aav87718 EST clone
36	161	20.2	697	10 ADE28664	Ades28664 Human NOV
37	161	20.2	697	12 ADM93409	Adm93409 Human NOV
38	142.5	17.8	1098	6 ABS70398	Abs70398 Human bon
39	98.5	12.3	4490	5 AAS76235	Aas76235 DNA encod
40	96.5	12.1	421	10 ACF57863	AcF57863 PLAC1 int
C 41	95	11.9	472	4 ABA25212	Abas25212 Probe #36
C 42	95	11.9	472	4 ABS28811	Abs28811 Human liv
C 43	88.5	11.1	413	4 AAI15997	Aai15997 Probe #59
C 44	88.5	11.1	413	4 ABA58401	Abas58401 Human foe
C 45	88.5	11.1	413	4 AAI38038	Aai38038 Probe #67

ALIGNMENTS

RESULT 1

AAH64743

ID AAH64743 standard; cDNA; 691 BP.

XX

AC AAH64743;

XX

DT 11-SEP-2001 (first entry)

XX Human secreted protein cDNA, SEQ ID NO: 19.

DE Human secreted protein; gene therapy; vaccine; treatment; diagnosis;

XX GENSET; ss.

KW Homo sapiens.

XX WO200142451-A2.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000WO-IB001938.

XX 08-DEC-1999; 99US-0169629P.

PR 06-MAR-2000; 2000US-0187470P.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX

DR WPI; 2001-367870/38.
XX P-PSDB; AAG89140.
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
PS Claim 7; Page 579; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with a patient's
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX
SQ Sequence 691 BP; 198 A; 126 C; 151 G; 216 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.52e-90 Length: 691
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 5 Gaps: 0
US-10-664-025-3903 (1-153) x AAH64743 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTTrpThrGlyAlaGluAsnLeu 20
DB 42 ATGGCGGTAGAAGTCTTGATGCTCTCGCTGCTTGTGATTGGACCGGTGCTGAGAACCTC 101
QY 21 HisValIysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 102 CATGTGAATAAATGAATGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGNA 161
QY 41 SerArgAsnLeuTyriLePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 162 AGCAGAAATCTGTATATATTTGGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAA 221
QY 61 ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr 80
DB 222 CGGATACATACATATGATATGAGTTTATATATCTTGTTCGTGATGTGGCATCAGGACA 281
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
DB 282 AGGGTAGTTCTTCAGGAACCTCTCCTTTTCAAACCGAGCTGTACTTACCCCAAGGAAT 341
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIlyssSerValTrp 120
DB 342 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 401
QY 121 LeuThrProValSerThrCluAsnGluIleIysLeuAspProSerProPheIleAlaAsp 140
DB 402 CTTACACCAAGTTTCTACTCAGAAATGAAATAAAATTTGGATCTAGTCCCTTTTATTCTGAC 461
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
DB 462 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 494

RESULT 2

ABK90053
ID ABK90053 standard; DNA; 1337 BP.
XX
AC ABK90053;
XX
DT 05-NOV-2002 (first entry)
XX
DE DNA encoding human GSSP3 polypeptide.
XX
KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
KW body mass; serum glucose regulation; body weight loss; obesity;
KW metabolic-related disorder; impaired glucose tolerance; stroke;
KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
KW hypertension; syndrome C; type I diabetes; type II diabetes;
KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
KW lipotrophy; physical performance; exercise; dyslexia; schizophrenia;
KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
KW gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 53..529
FT /tag= a
FT /product= "GSSP3 polypeptide"
FT sig_peptide 53..103
FT /tag= b
FT mat_peptide 104..526
FT /tag= c
XX
FN WO200260466-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-IB001333.
XX
PR 02-FEB-2001; 2001US-0266156P.
XX
PA (GEST) GENSET.
XX
PI Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
PI Bihain B;
XX
DR WPI; 2002-608487/65.
DR P-PSDB; ABG31324.
XX
PT Reducing circulating glucose levels or increasing insulin sensitivity,
PT useful for reducing body mass or preventing body weight gain, comprises
PT administering composition comprising GSSP3 polypeptide.
XX
PS Disclosure; Page 95-96; 97pp; English.
XX
CC The present invention relates to the isolation of human GSSP3
CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
CC polypeptide reduces circulating blood glucose levels, increases insulin
CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
CC polynucleotide sequences are useful in serum glucose regulation, fatty
CC acid metabolism, body weight loss, and prevention of body weight gain.
CC Compositions comprising GSSP3 polypeptides are useful for controlling
CC blood glucose levels, for treating metabolic-related diseases or
CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
CC syndrome C, type I or II diabetes, diabetes related complications,
CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
CC acanthosis nigricans, leprechaunism and lipotrophy). The polypeptides
CC are also useful to improve physical performance during work or exercise,
CC and to treat dyslexia, attention-deficit disorder, attention-
CC deficit/hyperactivity disorder, and psychiatric disorders such as
CC schizophrenia. The present sequence encodes human GSSP3 polypeptide
XX
SQ Sequence 1337 BP; 457 A; 214 C; 240 G; 426 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,93e-89 Length: 1337
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x ABK90053 (1-1337)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuThrGlyAlaGluAsnLeu 20
DB 53 ATGGCGTTAGAGTCTTGTATGCTCTCGCTGTCTTGTATTTGGACCGGTCTGAGAACCTC 112
QY 21 HisValValSerCysSerLeuAspTrpLeuMetValSerValLeuProValAlaGlu 40
DB 113 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 172
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAANTCTGATATATTTGGGATGAATATCATCTGGAAATGGCTGCCCTGCNAAT 232
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATGATATGATGTTTATATATCTTGTCTGATTTGGCATCAGGACA 292
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 293 AGGGTAGTTCTTCAGGAAACTCTCTTTTCAAAACGAGCTGTACTTATCCCAAGGAAT 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTyr 120
DB 353 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 121 LeuThrProValSerThrGluAsnGluIleLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACCACTTCTACTGAGATGAATAAATTTGGATCTAGTCTTATTTGCTGAC 472
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
DB 473 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 505

RESULT 3

ID ABS78646 standard; cDNA; 1480 BP.

XX AC ABS78646;

XX DT 16-DEC-2002 (first entry)

XX DE Human cDNA encoding, CGDD4, INCYTE 5284076CBI.

XX KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; porriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.

XX OS Homo sapiens.

XX WO200272830-A2.

XX PD 19-SEP-2002.

XX PF 08-FEB-2002; 2002WO-US003715.

XX 09-FEB-2001; 2001US-0268111P.
PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
XX 09-MAR-2001; 2001US-0274552P.
PA (INCY-) INCYTE GENOMICS INC.

XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Wallia NK;
DR WPI: 2002-723356/78.
XX P-PSDB; ABG97333.

XX New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.

XX Claim 5; Page 169; 181pp; English.

CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein

XX SQ Sequence 1480 BP; 463 A; 244 C; 265 G; 508 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,24e-89 Length: 1480
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x ABS78646 (1-1480)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuThrGlyAlaGluAsnLeu 20
DB 29 ATGGCGTTAGAGTCTTGTATGCTCTCGCTGTCTTGTATTTGGACCGGTCTGAGAACCTC 88
QY 21 HisValValSerCysSerLeuAspTrpLeuMetValSerValLeuProValAlaGlu 40
DB 89 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 148
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60

```

Db      149 AGCAGAAATCTGTATATATTTGCGATGAATTTACATCTGGGAATGGCTGCCCTGCAAAAT 208
QY      61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db      209 CGGATACATACATATGATATGAGTTTATATATCTTTGTCGATTTGGCATCAGGACA 268
QY      81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db      269 AGGGTAGTTTCTGAGGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 328
QY      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr 120
Db      329 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 388
QY      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db      389 CTTACACCACTTCTACTGAGAAATGAAATGAAATTTGGATCTTAGTCTTTTATTGCTGAC 448
QY      141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db      449 TTTACAGACACAGCAGAGAGTTAGGATTATTA 481

RESULT 4
ID AAC69528 standard; DNA; 1898 BP.
XX AAC69528;
AC AAC69528;
XX
DT 31-JAN-2001 (first entry)
XX
DE Human secreted protein gene 17 clone HTELS08.
XX
KW Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
KW nootropic; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; gene therapy; infection; human; secreted protein; ss.
XX
OS Homo sapiens.
XX
PN WO200061623-A1.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US008979.
XX
PR 09-APR-1999; 99US-0128693P.
PR 26-APR-1999; 99US-0130991P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ni J, Konatsoulis GA, Rosen CA, Soppet DR, Shi Y;
PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;
PI Young PE;
XX
DR WPI; 2000-647418/62.
XX
PT New nucleic acid molecules encoding 62 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 1; Page 543-544; 716pp; English.
XX
CC The invention relates to the isolation of genes (AAC69512-C69587)
CC encoding 62 human secreted proteins (AAB38321-B38396). The genes can be
CC used to generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (AAC69503) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.

```

```

CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of
CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;
CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angiogenesis;
CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections
CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.
CC corneal infection. The polypeptides can also be used to aid wound healing
CC and epithelial cell proliferation, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis
XX
SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:
  Pred. No.:      3.19e-89      Length:      1898
  Score:          789.00        Matches:      151
  Percent Similarity: 100.0%    Conservative: 0
  Best Local Similarity: 100.0%  Mismatches: 0
  Query Match:      98.7%       Indels:      0
  DB:                3          Gaps:        0

US-10-664-025-3903 (1-153) x AAC69528 (1-1898)
QY      1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTyrThrGlyAlaGluAsnLeu 20
Db      15 ATGGCGTTAGAAATCTTGTATCTCTCGCTCTGTCTTGTATTTGGACCGGTGAGAACCTC 74
QY      21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db      75 CATGTGAAAAATAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTCCAGAA 134
QY      41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db      135 AGCAGAAATCTGTATATATTTTGGGATGAAATTACATCTGGGAATGGGCTGCCCTGCAAAAT 194
QY      61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db      195 CGGTACATACATATGATATATGAGTTTATATATCTTGTTCGTGATTTGGGATCAGGACA 254
QY      81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db      255 AGGTAGTGTTCGAGGAAACTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 314
QY      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr 120
Db      315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 374
QY      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db      375 CTTACACCACTTCTACTGAGAAATGAAATGAAATTTGGATCTCTAGTCTTTTATTGCTGAC 434
QY      141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db      435 TTTACAGACACAGCAGAGAGTTAGGATTATTA 467

RESULT 5
ID AAC69528 standard; cDNA; 1898 BP.
XX AAC69528;
AC AAC69528;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein coding sequence, SEQ ID 309.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
KW vulnary; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200295010-A2.
XX
PD 28-NOV-2002.

```

XX PF 19-MAR-2002; 2002WO-US009785.
 XX PR 21-MAR-2001; 2001US-0277340P.
 XX PR 19-JUL-2001; 2001US-0306171P.
 XX PR 13-NOV-2001; 2001US-0331287P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Ruben SM;
 XX DR WPI; 2003-129429/12.
 XX PT Novel human secreted proteins, useful for detecting, preventing,
 XX PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
 XX PT disorders such as arrhythmia.
 XX PS Claim 21; SEQ ID NO 309; 1881pp; English.
 XX
 CC The present invention relates to novel human secreted proteins (ABR47633-
 CC ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
 CC and their coding sequences are useful for the preparation of a diagnostic
 CC or pharmaceutical composition for diagnosing or treating a cardiovascular
 CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
 CC arteriosclerosis and myocardial ischaemia), neural disorders, immune
 CC system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,
 CC proliferative disorders and/or cancerous diseases and conditions, for
 CC wound healing and epithelial cell proliferation, to treat inflammation or
 CC infection, for treating thrombosis and arteriosclerosis, for treating or
 CC preventing neural damage which occurs in neuronal disorders or
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
 CC disease, to enhance bone and periodontal regeneration and aid in tissue
 CC transplants or bone grafts, to prevent skin aging or hair loss, to
 CC stimulate growth and differentiation of haematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines, to maintain
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues, to increase or decrease differentiation or proliferation of
 CC embryonic stem cells, or to modulate mammalian characteristics or
 CC metabolism. Note: The sequence data for this patent was published in
 CC electronic format and is available from WIPO at
 CC ftp.wipo.int/pub/published_pt_sequences
 XX SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:
 Pred. No.: 3.19e-89 Length: 1898
 Score: 789.00 Matches: 151
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 98.7% Indels: 0
 DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x ACC50642 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuTTrpThrClyAlaGluAsnLeu 20
 DB 15 ATGGCGGTTAGAGTCTTGATGCTCTCGCTGCTTGAATGACCGGTCTGAGAACCTC 74
 QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
 DB 75 CATGTGAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGAGAA 134
 QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
 DB 135 AGCAGAAATCTGATATATTTTCGGATGAATTACATCTCGGAATGGCTGCCCTGCMAAT 194
 QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
 DB 195 CGGATACATACATATGATATGATGATATATATCTTGTTCGTGATGTGGCATCAGGACA 254
 QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100

Db 255 AGGTTAGTTTCTGAGGAACCTCTCTTTTCAACCCGAGCTGTACTTTACCCCAAGNAAT 314
 QY 101 IleAspHisAspProGlnGluLeuHisLeuGluCysSerThrSerArgIysSerValTyr 120
 Db 315 ATAGATCATGACCTCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 374
 QY 121 LeuThrProValSerThrGluAsnGluIleLeuLeuAspProSerProPheIleAlaAsp 140
 Db 375 CTTACACCAAGTTTCTACTGAGAAATGAATAAATTGGATCTAGTCTCTTTATTGCTGAC 434
 QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
 Db 435 TTTCAGACAACAGCAGACAGTATAGGATTATTA 467
 RESULT 6
 ID ABZ71354 standard; cDNA; 1898 BP.
 XX AC ABZ71354;
 XX DT 04-APR-2003 (first entry)
 XX DE Secreted protein-encoding gene 165 cDNA clone HTLS08, SEQ ID NO:175.
 XX KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
 KW immune disorder; inflammation; infection; wound healing; drug screening;
 KW chromosome identification; chromosome mapping; cytostatic;
 KW antiinflammatory; immunosuppressive; vulnery; gene therapy; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200276488-A1.
 XX PD 03-OCT-2002.
 XX PF 19-MAR-2002; 2002WO-US008276.
 XX PR 21-MAR-2001; 2001US-0277340P.
 XX PR 19-JUL-2001; 2001US-0306171P.
 XX PR 13-NOV-2001; 2001US-0331287P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Ruben SM;
 XX DR WPI; 2003-029900/02.
 XX DR P-PSDB; ABR00175.
 XX PT New human secreted proteins and nucleic acids, useful for detecting,
 XX PT preventing, diagnosing, prognosticating, treating and/or ameliorating
 XX PT e.g. gastrointestinal diseases and disorders, or cancers.
 XX PS Claim 21; Page 853-854; 1216pp; English.

XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening, and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing, treating, ameliorating or preventing
 CC digestive disorders. Such conditions include disorders of the mouth,
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary
 CC tract and pancreas, and include cancers of these organs and tissues. The
 CC secreted proteins and their nucleic acids may also be used in the
 CC treatment of immune disorders, inflammation, infection,
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
 CC of the invention may be used for chromosome identification, chromosome
 CC mapping, in gene therapy, for identifying individuals from minute

CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein-
CC encoding cDNA clone of the invention

XX
SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3.19e-89 Length: 1898
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x AB271354 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeu 20
DB 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGTGCTGAGAACCTC 74
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 75 CATGTGAAATAAGTTGCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 134
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 135 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTCGGAATGGGCTGCCCTGCAAT 194
QY 61 ArgIleHisThrTyrlleValTyrllePheLeuLeuValArgAspCysGlyIleArgThr 80
DB 195 CGGATACATACATATGATATAGTTTATATATCTTTGTTGATTTGGCATCAGGACA 254
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
DB 255 AGGTTAGTTTCTGAGGAAACTCTCCTTTTCAACCGAGCTGACTTACCCCAAGGAAT 314
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
DB 315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTTGG 374
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 375 CTTACACCACTTCTACTCAGAAATGAAATGAAATTTGGATCCTAGTCTTTTATTGCTGAC 434
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
DB 435 TTTCAGACAAACAGCAGAGATTAGGATTATTA 467

RESULT 7

ID ADB91291
XX ADB91291 standard; cDNA; 1898 BP.
XX ADB91291;
AC ADB91291;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human secreted protein cDNA #SEQ ID 237.
XX
KW Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.
XX
OS Homo sapiens.
XX
XX WO2003004622-A2.
XX
XX 16-JAN-2003.
XX
XX 19-MAR-2002; 2002WO-US008124.
XX
XX 21-MAR-2001; 2001US-0277340P.
XX
XX 19-JUL-2001; 2001US-0306171P.
XX
XX 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;
XX WPI; 2003-229407/22.
XX
XX Nucleic acid encoding a human secreted protein is useful in diagnosing or
XX treating diabetes or conditions related to diabetes.
XX
XX Claim 9; SEQ ID NO 237; 1537pp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-
XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX ADB91834. Also disclosed is a recombinant vector comprising a
XX polynucleotide of the invention, and a recombinant host cell comprising
XX the recombinant vector. The polypeptide of the invention is useful in
XX identifying a binding partner by contacting the polypeptide with a
XX binding partner, and determining whether the binding partner increases or
XX decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX antibody or its fragment, agonist or antagonist are useful for preparing
XX a pharmaceutical composition for diagnosing or treating diabetes or
XX conditions related to diabetes. The present sequence is that of the human
XX immunoglobulin FC portion used to generate fusion proteins, increasing
XX the stability of the fused protein as compared to the secreted protein
XX only. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3.19e-89 Length: 1898
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-3903 (1-153) x ADB91291 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeu 20
DB 15 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGTGCTGAGAACCTC 74
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 75 CATGTGAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 134
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 135 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTCGGAATGGGCTGCCCTGCAAT 194
QY 61 ArgIleHisThrTyrlleValTyrllePheLeuLeuValArgAspCysGlyIleArgThr 80
DB 195 CGGATACATACATATGATATAGTTTATATATCTTTGTTGATTTGGCATCAGGACA 254
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
DB 255 AGGTTAGTTTCTGAGGAAACTCTCCTTTTCAACCGAGCTGACTTACCCCAAGGAAT 314
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
DB 315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTTGG 374
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 375 CTTACACCACTTCTACTCAGAAATGAAATGAAATTTGGATCCTAGTCTTTTATTGCTGAC 434
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
DB 435 TTTCAGACAAACAGCAGAGATTAGGATTATTA 467

RESULT 8

ADC73723

ID	ADC73723 standard; DNA; 1898 BP.	
XX	ADC73723;	
AC	01-JAN-2004 (first entry)	
XX	Human secreted protein-related DNA - SEQ ID 356.	
DE	antianaemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;	
XX	antidiabetic; immunosuppressive; dermatological; nephrotropic;	
KW	antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;	
KW	fungicide; antiparasitic; antiarteriosclerotic; vulnery; cytostatic;	
KW	haemopoietic; haematologic; anaemia; autoimmune disorder;	
KW	rheumatoid arthritis; inflammation; Grave's disease; diabetes;	
KW	systemic lupus erythematosus; glomerulonephritis; neurodegenerative;	
KW	Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;	
KW	cancer; bacterial; viral; fungal; parasitic infection; gene therapy;	
KW	human; gene; ds.	
XX	Homo sapiens.	
XX	WO2003038063-A2.	
PN	08-MAY-2003.	
XX	19-MAR-2002; 2002WO-US008277.	
XX	21-MAR-2001; 2001US-0277340P.	
PR	19-JUL-2001; 2001US-0306171P.	
PR	13-NOV-2001; 2001US-0331287P.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Ruben SM;	
PI	WPI; 2003-430516/40.	
XX	P-PSDB; ADC74338.	
DR	New human secreted polypeptide for diagnosing, preventing or treating	
XX	hematopoietic or hematologic disorders (e.g. anemia), autoimmune	
PT	disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or	
PT	atherosclerosis).	
XX	Claim 27; SEQ ID NO 356; 2272pp; English.	
PS	The invention relates to a novel human secreted polypeptide comprising a	
CC	defined sequence given in the specification. The polypeptide, nucleic	
CC	acid molecule, antibody, agonist or antagonist of the invention may be	
CC	useful for preparing a composition for diagnosing or treating a	
CC	haemopoietic or haematologic disorder such as anaemia, autoimmune	
CC	disorders such as rheumatoid arthritis, inflammation, Grave's disease,	
CC	diabetes, systemic lupus erythematosus or glomerulonephritis,	
CC	neurodegenerative disorders including Parkinson's disease and Alzheimer's	
CC	disease, wounds and hyperproliferative disorders including	
CC	atherosclerosis or cancer, as well as bacterial, viral, fungal or	
CC	parasitic infections. The polypeptide may also be used during gene	
CC	therapy procedures and for identifying a binding partner by contacting	
CC	the polypeptide with a binding partner and determining whether the	
CC	binding partner increases or decreases the activity of the polypeptide.	
CC	The current sequence is that of the human secreted protein-related DNA of	
CC	the invention.	
XX	Sequence 1898 BP; 595 A; 315 C; 332 G; 654 T; 0 U; 2 Other;	
SQ	Alignment Scores:	
	Pred. No.: 3,19e-89 Length: 1898	
Score:	789.00 Matches: 151	
Percent Similarity:	100.0% Conservative: 0	
Best Local Similarity:	100.0% Mismatches: 0	
Query Match:	98.7% Indels: 0	
DB:	10 Gaps: 0	
DR	US-10-664-025-3903 (1-153) x ADC73723 (1-1898)	

QY	1	MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuLeuTrpThrGlyAlaGluAsnLeu	20
DB	15	ATGCGGTAGAGTCTTGATGCTCCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC	74
QY	21	HisValIleSerCysSerLeuAspTyrLeuMetValSerValIleProValAlaGlu	40
DB	75	CATGTGAATAAGTTGCTCTCTGGACTGGTGTGATGCTCAGTTATCCCAAGTTCGAA	134
QY	41	SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn	60
DB	135	AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGCTGCCCTGCAAT	194
QY	61	ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr	80
DB	195	CGGATACATACATATGATATATATATATATCTTTGTTGTTGTTGTCATCAGGACA	254
QY	81	ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn	100
DB	255	AGGTTAGTTCTGAGGAAATCTCTCTTTTCAACCGAGCTGACTTTACCCCAAGAAAT	314
QY	101	IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr	120
DB	315	ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTG	374
QY	121	LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp	140
DB	375	CTTACACCACTTCTACTGAGAATGAATAAAATTGGATCTAGTCTCTTTTATTGCTGAC	434
QY	141	PheGlnThrThrAlaGluGluLeuGlyLeuLeu	151
DB	435	TTTCAGACCAACAGCAGAGATTAGGATTATTA	467
RESULT	9		
ID	ADB63550		
ID	ADB63550 standard; cDNA; 1492 BP.		
XX	AC	ADB63550;	
XX	04-DEC-2003 (first entry)		
DT	Human cDNA encoding clone TESTI20282420.		
DE	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;		
XX	tissue regeneration; cell regeneration; membrane protein;		
KW	signal transduction-related protein; transcription-related protein;		
KW	osteoporosis; neurological disease; cancer; tumour.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	CDS	52..528	
FT		/*tag= a	
FT		/product= "Clone TESTI20282420 protein"	
XX	EP1308459-A2.		
PN	07-MAY-2003.		
XX	28-MAR-2002; 2002EP-00007401.		
XX	05-NOV-2001; 2001JP-00379298.		
PR	25-JAN-2002; 2002US-00350978.		
XX	(HELI-) HELIX RES INST.		
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.		
XX	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;		
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;		
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;		
XX	WPI; 2003-450961/43.		
DR	P-PSDB; ADB65520.		

XX New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page; 222pp; English.
XX
CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesising the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX
SQ Sequence 1492 BP; 452 A; 250 C; 272 G; 518 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.26e-89 Length: 1492
Score: 785.00 Matches: 150
Percent Similarity: 99.3% Conservativity: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 98.2% Indels: 0
DB: 10 Gaps: 0

US-10-664-025-3903 (1-153) x ADB63550 (1-1492)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeu 20
Db 52 ATGGCGTTAGAAAGCTTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 111
Qy 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 112 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGCCCTCAGTTATCCAGTTGCAGAA 171
Qy 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 172 AGCAGAAATCTGTATATATTTTCGGATGAATATACATCTGGGAATGGCTGCCCTCAAT 231
Qy 61 ArgIleHisThrTyrrValTrpGluPheIleTyrlleValAlaArgCysGlyIleArgThr 80
Db 232 CGGATACATACATATGATATGAGTTTATATATCTTGTTCGGATTTGGCATCAGGACA 291
Qy 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
Db 292 AGGGTAGTTTCTGAGGAAATCTCTCTTTTCAACCGAGCTGTACTTACCCCAAGGAAT 351
Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Db 352 ATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAAGTGTG 411
Qy 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db 412 CTTACACAGTTTCTACTGAGAAATGAATATAAATTGGATCCTAGTCCITTTATTTGCTGAC 471
Qy 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151

Db 472 TTTCAGACACAGCAGAGAGAGTTAGGATTATTA 504
RESULT 10
ABZ36691
ID ABZ36691 standard; cDNA; 497 BP.
XX
XX AC ABZ36691;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET coding sequence SEQ ID 564.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiac;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
XX 24-OCT-2002.
XX
XX 18-APR-2001; 2001WO-IB000914.
XX
PR 18-APR-2001; 2001WO-IB000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX
XX WPI; 2003-075548/07.
XX
PT New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
PS Claim 12; Page 579; 735pp; English.
XX
CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
CC polynucleotides are useful for constructing or expanding chromosome maps
XX
SQ Sequence 497 BP; 127 A; 99 C; 117 G; 154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.78e-88 Length: 497
Score: 775.00 Matches: 148
Percent Similarity: 100.0% Conservativity: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.0% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x ABZ36691 (1-497)

Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeu 20
Db 52 ATGGCGTTAGAAAGCTTTGATGCTCTCGCTGCTTGTGATTTGACCGGTGCTGAGAACCTC 111
Qy 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 112 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGCTCAGTTATCCAGTTGCAGAA 171
Qy 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 172 AGCAGAAATCTGTATATATTTTCGGATGAATATACATCTGGGAATGGCTGCCCTCAAT 231

QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
 DB 232 CGGATACATACATATGATATAGTTTATATATCTTGTTCGTATGTCGATCAGGACA 291
 QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
 DB 292 AGGGTAGTTCTCGAGGAACCTCTCTTTTCAACCGAGCTGTACTTATCCCAAGGAAT 351
 QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr 120
 DB 352 ATAGATCATACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 411
 QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
 DB 412 CTTACACCACTTCTACTGAGATGAATAAATTTGGATCCTAGTCTTTATTTGCTGAC 471
 QY 141 PheGlnThrThrAlaGluGluLeu 148
 DB 472 TTTTCAGACACAGCAGGAAGATTGA 495

RESULT 11

ABZ36690
 ID ABZ36690 standard; cDNA; 500 BP.

XX AC ABZ36690;

XX DT 21-FEB-2003 (first entry)

XX DE Human GENSET coding sequence SEQ ID 563.

XX KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;
 KW central nervous system; cardiovascular; gastrointestinal; gene; ss.

XX OS Homo sapiens.

XX PN WO200283898-A1.

XX PD 24-OCT-2002.

XX PF 18-APR-2001; 2001WO-IB000914.

XX PR 18-APR-2001; 2001WO-IB000914.

XX PA (GEST) GENSET.

XX PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;

XX DR WPI; 2003-075548/07.

XX PT New GENSET polynucleotides and polypeptides, useful for treating heavy
 PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
 PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
 PT toxicity.

XX PS Claim 12; Page 578; 735pp; English.

XX CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
 CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
 CC and polypeptides are useful in screening and diagnostic assays for
 CC abnormal GENSET expression and/or biological activity. They are also
 CC useful for screening of compounds for treating or preventing GENSET-
 CC related disorders, such as heavy metal toxicity, cancer, inflammatory
 CC diseases, immune disorders, and the neuromuscular, central nervous system
 CC (CNS), cardiovascular or gastrointestinal effects of the toxicity. The
 CC polynucleotides are useful for constructing or expanding chromosome maps

XX SQ Sequence 500 BP; 129 A; 100 C; 113 G; 155 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 2.29e-85 Length: 500

Score: 752.00 Matches: 143
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 94.1% Indels: 0
 DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x ABZ36690 (1-500)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTyrThrGlyAlaGluAAsnLeu 20
 DB 72 ATGGCGCTTGAAGTCTTGATGCTCCTCGCTGTCTTCATTGGACCGGTGCTGAGAACCTC 131
 QY 21 HisValLysIleSerCysSerLeuAspTyrLeuMetValSerValIleProValAlaGlu 40
 DB 132 CATGTGAATAAATGAATGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 191
 QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
 DB 192 AGCAGAAATCTGATATATATTTGGGATGAATTACATCTGGGAATGGCTGCCCTGCCAAT 251
 QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
 DB 252 CGGATACATACATATGATATGAGTTTATATATCTTGTTCGTATGTCGATGTCGATCAGGACA 311
 QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
 DB 312 AGGGTAGTTTCTGAGGAACCTCTCTCTTTTCAACCGAGCTGTACTTTATCCCAAGGAAT 371
 QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr 120
 DB 372 ATAGATCATGACCCCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAATCAGTGTGG 431
 QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
 DB 432 CTTACACCACTTCTACTGAGAAATGAATAAAATTTGGATCCTAGTCTTTATTTGCTGAC 491
 QY 141 PheGlnThr 143
 DB 492 TTTTCAGACA 500

RESULT 12

AAH64748

ID AAH64748 standard; cDNA; 470 BP.

XX AC AAH64748;

XX DT 11-SEP-2001 (first entry)

XX DE Human secreted protein cDNA, SEQ ID NO: 24.

XX KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KW GENSET; ss.

XX OS Homo sapiens.

XX PN WO200142451-A2.

XX PD 14-JUN-2001.

XX PF 07-DEC-2000; 2000WO-IB001938.

XX PR 08-DEC-1999; 99US-0169629P.

XX PR 06-MAR-2000; 2000US-0187470P.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX DR WPI; 2001-367870/38.

XX DR P-PSDB; AAG89145.

XX PT Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of

PT diseases, and for diagnosis of those diseases.

PS Claim 7; Page 583; 921pp; English.

XX
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patients own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention

XX SQ Sequence 470 BP; 137 A; 87 C; 107 G; 139 T; 0 U; 0 Other;

Alignment Scores: 1,66e-44 Length: 470
Pred. No.: 429.00 Matches: 82
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 53.7% Gaps: 0
DB:

US-10-664-025-3903 (1-153) x AAH64748 (1-470)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuIleTrpThrGlyAlaGluAsnLeu 20
Db 137 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGTGACCGGTCTGAGAACCTC 196
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 197 CATGTGAATAAGTTGCTCTGGACTGTTGATGGTCTCATGTTATCCAGTTGCAGAA 256
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 257 ACGAGAAATCTGATATATTTGCGATGATTTACATCTGGGATGGCTGCCCTGCAAT 316
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 317 CGGATACATACATATGATATGATGATTTATATATCTTTGTTGTTGTCGATTCAGGACA 376
QY 81 ArgVal 82
Db 377 AGGGTA 382

RESULT 13

ADJ12678

ID ADJ12678 standard; DNA; 5013 BP.

XX

AC ADJ12678;

XX 20-MAY-2004 (first entry)

DE DNA fragment of a BAC clone that encodes a human secreted protein Seq532.

XX human; secreted; cancer; haematopoietic disease; anaemia;
KW multiple myeloma; reproductive system disorder; prostatitis;
KW inguinal hernia; musculoskeletal disease; systemic lupus erythematosus;
KW gout; cardiovascular disease; arrhythmia; hypernatraemia; fetal disease;
KW fetal alcohol syndrome; Down's syndrome; excretory disease;
KW urinary incontinence; renal disorder; neural; sensory disease;
KW Alzheimer's disease; meningitis; respiratory disease; emphysema;
KW occupational lung disease; endocrine disease; diabetes;

KW glomerulonephritis; digestive disease; portal hypertension;
KW irritable bowel syndrome; epithelial disease; scleroderma;
KW epidermolysis bullosa; cytostatic; antianemic; antiarthritic;
KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
KW antiparasitic; antibacterial; osteopathic; dermatological; antigout;
KW immunomodulator; antiarrhythmic; cardiac; nootropic; antileptic;
KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
XX antidiabetic; anabolic; hypertensive; vulnery; ds.

XX Homo sapiens.

OS US2004010132-A1.

XX 15-JAN-2004.

XX 30-OCT-2001; 2001US-00984429.

XX 09-OCT-1997; 97US-0061463P.

XX 09-OCT-1997; 97US-0061527P.

XX 09-OCT-1997; 97US-0061529P.

XX 09-OCT-1997; 97US-0061532P.

XX 09-OCT-1997; 97US-0061538P.

XX 09-OCT-1997; 97US-0071498P.

XX 08-OCT-1998; 98WO-US021142.

XX 08-APR-1999; 99US-00288143.

XX 01-NOV-2000; 2000US-0244591P.

XX (ROSE/) ROSEN C A.

XX (BREW/) BREWER L A.

XX (DUAN/) DUAN R D.

XX (RUBE/) RUBEN S M.

XX (FLOR/) FLORENCE K A.

XX (GREE/) GREENE J M.

XX (YOUN/) YOUNG P E.

XX (FERR/) FERRIE A M.

XX (YUGG/) YU G.

XX (FLOR/) FLORENCE C.

XX (EBNE/) EBNER R.

XX (OLSE/) OLSEN H.

XX Rosen CA, Brewer LA, Duan RD, Ruben SM, Florence KA, Greene JM;

PI Young PE, Ferrie AM, Yu G, Florence C, Ebner R, Olsen H;

XX WPI; 2004-090518/09.

XX New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating, preventing or ameliorating diseases or disorders e.g. cancer,
PT anemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's
PT disease.

XX Disclosure; SEQ ID NO 532; 286pp; English.

XX This invention relates to novel polynucleotides encoding human secreted
CC proteins. Specifically, it refers to the vectors, host cells, recombinant
CC and synthetic methods for producing human polynucleotides, polypeptides
CC and antibodies. Furthermore, it relates to screening methods to identify
CC agonists and antagonists that can be used to inhibit or enhance the
CC production and function of the secreted proteins. The present invention
CC describes these compositions as useful for diagnosing, treating or
CC preventing disorders such as cancer, haematopoietic diseases including
CC anaemia and multiple myeloma, reproductive system disorders including
CC prostatitis and inguinal hernia, musculoskeletal diseases including
CC systemic lupus erythematosus and gout, cardiovascular disease including
CC arrhythmia and hypernatraemia, mixed fetal diseases including fetal
CC alcohol syndrome and Down's syndrome, excretory diseases including
CC urinary incontinence and renal disorders, neural or sensory disease
CC including Alzheimer's disease and meningitis, respiratory disease
CC including emphysema and occupational lung disease, endocrine diseases
CC including diabetes and glomerulonephritis, digestive diseases including
CC portal hypertension and irritable bowel syndrome and connective tissue or
CC epithelial diseases including scleroderma and epidermolysis bullosa. As
CC such, there are various activities such as cytostatic, antianemic,
CC antiarthritic, antiasthmatic, anti-HIV, immunosuppressive,

CC antiinflammatory, antipsoriatic, antibacterial, osteopathic,
 CC dermatological, antitumor, immunomodulator, antiarrhythmic, cardiant,
 CC neotropic, antilipemic, nephrotropic, uropathic, neuroprotective,
 CC antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and
 CC vulnerable. This polynucleotide is a DNA fragment of a BAC clone that
 CC encodes a human secreted protein of the invention. NOTE: This sequence
 CC does not appear in the printed specification but has been obtained in
 CC electronic format from the US patent office at the following web site
 CC www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.
 XX

SQ Sequence 5013 BP; 1442 A; 909 C; 826 G; 1836 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6,29e-30 Length: 5013
 Score: 325.50 Matches: 95
 Percent Similarity: 21.8% Conservative: 0
 Best Local Similarity: 21.8% Mismatches: 0
 Query Match: 40.7% Indels: 341
 DB: 12 Gaps: 1

US-10-664-025-3903 (1-153) x ADJ12678 (1-5013)

Qy 22 ValylSerCysSerLeuAspTrpLeuMetValSerValIleProValaIaGluSer 41
 Db 1 GTGAAATAAGTTGCTCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAAAGC 60
 Qy 42 ArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArg 61
 Db 61 AGAAATCTGTATATATTGCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAAATCGG 120
 Qy 62 IleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArg 81
 Db 121 ATACATACATGATGATATAGATTATATATCTTGTTCGTGATTTGGCATCAGCAAGG 180
 Qy 81 ----- 81
 Db 181 GTAAGAACAGTGATTGTTGTAAATAAATACTGCATGTTTGTGCAGACTTTATGCTAGT 240
 Qy 81 ----- 81
 Db 241 ATTTAAATTTGCTTTTAAATAAATAATGTTTGTCTTCTCATTTAGTAAATTTATGGCAGG 300
 Qy 81 ----- 81
 Db 301 GGCAACGGAAATTATCATTTATTTTCACTGAATATTTGGTCAATCTGTGTGCACAGGAC 360
 Qy 81 ----- 81
 Db 361 CTTAGATCACAGATTTTAAATTTGGCTAGGGAATGATTTCAAGTTGCCCTTAGGCTACT 420
 Qy 81 ----- 81
 Db 421 ATTATATTTTACTAAATATTAGCAGACTGCTTTTAAATTTTGTGTGATTTCTTAAATGTGG 480
 Qy 81 ----- 81
 Db 481 AAGTCAACCAATCTTGGGAAGCTCCATCATGTTCAATGTTAGTCAATTTTGAATCTATCC 540
 Qy 81 ----- 81
 Db 541 ATGTTCCATGAGGATTTGTGTAATATATATAGTGGATTTTTCATTGAAATGAAATCTTTGT 600
 Qy 81 ----- 81
 Db 601 CATTCAGTAGGCACAAATCTAAATTTGGTTTATCTTCTGGGATATTATCTTCTCACAG 660
 Qy 81 ----- 81
 Db 661 ATTCCAAATTTAAGCAATTTATATTCACTACACAAATATTTTGGGCCAGTGACCAAT 720
 Qy 81 ----- 81
 Db 721 GATAATAACAATCATTTTGAATAAAGTATTTTATTTTAGAAATTTGGGAACCTGGAGCTT 780

Qy 81 ----- 81
 Db 781 CAGAAGGTGCTTTTACCAGCTTTTCCAGGATGCACTGCTTTGTTGGTGGTAGTTTAGATAC 840
 Qy 81 ----- 81
 Db 841 CAAACCCAGGATGAATACTCTCCGAAGCTTGTGCTTTTAAACCACTCTCCGTTATTGACTT 900
 Qy 81 ----- 81
 Db 901 ACAATCACCTTCTCAAAGGAACCTGTTGGAACCTTTTTCATTAAATGAATCAATAAGATG 960
 Qy 81 ----- 81
 Db 961 GCTGTAAGCGGCAGTTTAAATTTTGTGCAATGAATTAAGATGTTTGTGCCCCCAGGCATT 1020
 Qy 81 ----- 81
 Db 1021 TTAATTCACCTCTTAGGATAACACTCTCTTCTGTGCCAGTGAGTTGTCATATCTTTAT 1080
 Qy 81 ----- 81
 Db 1081 TTATATTTAAGATGACAGTAGCCATTTTGTGAGATCATGAATGTTTTTATACCTTATTC 1140
 Qy 81 ----- 81
 Db 1141 AAGTGTGTGTTTCTCATCTGTTACTTCCAAATATCTTCCACTGAAACTTTCATGCTCTCC 1200
 Qy 82 ---ValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
 Db 1201 TAGGTAGTTTCTGAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 1260
 Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArg 116
 Db 1261 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACACTCTAGG 1308
 RESULT 14
 ABK90052
 ID ABK90052 standard; DNA; 10115 BP.
 XX
 AC ABK90052;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human GSSP3 genomic sequence.
 XX
 KW Human; GSSP3; circulating blood glucose level; insulin sensitivity;
 KW body mass; serum glucose regulation; body weight loss; obesity;
 KW metabolic-related disorder; impaired glucose tolerance; stroke;
 KW insulin resistance; hyperlipidemia; atherosclerosis; heart disease;
 KW hypertension; syndrome C; type I diabetes; type II diabetes;
 KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
 KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;
 KW lipodystrophy; physical performance; exercise; dyslexia; schizophrenia;
 KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder;
 KW gene; db.
 XX
 OS Homo sapiens.
 XX
 PN WO200260466-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 01-FEB-2002; 2002WO-IB001333.
 XX
 PR 02-FEB-2001; 2001US-0266156P.
 XX
 PA (GEST) GENSET.
 XX
 PI Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;
 PI Bihain B;
 XX

DR WPI; 2002-608487/65.
XX Reducing circulating glucose levels or increasing insulin sensitivity,
PT useful for reducing body mass or preventing body weight gain, comprises
PT administering composition comprising GSSP3 polypeptide.
XX
XX Disclosure; Page 90-95; 97pp; English.
XX
XX The present invention relates to the isolation of human GSSP3
CC polypeptide, and polynucleotide sequences that encode it. The GSSP3
CC polypeptide reduces circulating blood glucose levels, increases insulin
CC sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
CC polynucleotide sequences are useful in serum glucose regulation, fatty
CC acid metabolism, body weight loss, and prevention of body weight gain.
CC Compositions comprising GSSP3 polypeptides are useful for controlling
CC blood glucose levels, for treating metabolic-related diseases or
CC disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
CC hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
CC syndrome C, type I or II diabetes, diabetes related complications,
CC microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
CC syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
CC acanthosis nigricans, leprechaunism and lipodystrophy). The polypeptides
CC are also useful to improve physical performance during work or exercise,
CC and to treat dyslexia, attention-deficit disorder, attention-
CC deficit/hyperactivity disorder, and psychiatric disorders such as
CC schizophrenia. The present sequence encodes human GSSP3 polypeptide
XX
SQ Sequence 10115 BP; 2860 A; 1941 C; 1815 G; 3475 T; 0 U; 24 Other;

Alignment Scores:
Pred. No.: 1.99e-29 Length: 10115
Score: 325.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 40.7% Indels: 0
DB: 6 Gaps: 0

US-10-664-025-3903 (1-153) x ABK90052 (1-10115)

QY 22 VallysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValalaGluSer 41
Db 5444 GTGAATAAAGTTGCTCTCGACATGGTTGATGGTCTCAGTTATCCCGATTCGACAAAGC 5503

QY 42 ArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArg 61
Db 5504 AGAATCTGTATATATTTCGGATGATTAATACATCTGGGATGGCTGCCCTGCANAATCG 5563

QY 62 IleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArg 81
Db 5564 ATACATACATATGATATACGTTTATATATCTTGTTCGTGATTTGGCATCAGGACAAG 5623

QY 82 Val 82
Db 5624 GTA 5626

RESULT 15
ADE28662
ID ADE28662 standard; cDNA; 838 BP.
AC ADE28662;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human NOV12a cDNA - SEQ ID 39.
XX
XX NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antiinflammatory; dermatological;
KW antiasthmatic; antilipaeamic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;

KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; ss; gene; NOV.
XX Homo sapiens.
XX WO2003040330-A2.
XX
XX 15-MAY-2003.
XX
XX 05-NOV-2002; 2002WO-US035536.
XX
XX 05-NOV-2001; 2001US-0338626P.
XX 05-DEC-2001; 2001US-0336600P.
XX 07-DEC-2001; 2001US-0338285P.
XX 12-DEC-2001; 2001US-0341348P.
XX 17-DEC-2001; 2001US-0341477P.
XX 17-DEC-2001; 2001US-0341540P.
XX 20-DEC-2001; 2001US-0342592P.
XX 27-DEC-2001; 2001US-0344297P.
XX 31-DEC-2001; 2001US-0344903P.
XX 17-APR-2002; 2002US-0373289P.
XX 15-MAY-2002; 2002US-0380981P.
XX 17-MAY-2002; 2002US-0381495P.
XX 28-MAY-2002; 2002US-0383534P.
XX 28-MAY-2002; 2002US-0383744P.
XX 29-MAY-2002; 2002US-0383829P.
XX 29-MAY-2002; 2002US-0384024P.
XX 07-AUG-2002; 2002US-0401788P.
XX 26-AUG-2002; 2002US-0406353P.
XX 31-OCT-2002; 2002US-00287971.
XX (CURA-) CURAGEN CORP.
XX
XX Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL; Eisen A;
PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K;
PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI Pena CRA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA;
PI Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VI, Twomlow N;
PI Vernet CAM, Zerhusen BD, Zhong M;
XX
XX WPI: 2003-441555/41.
XX P-PSDB; ADE28663.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
XX Claim 20; SEQ ID NO 39; 447pp; English.
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC dermatological, antiasthmatic and antilipaeamic activities. The
CC polypeptides, nucleic acid molecules and antibodies may be useful for
CC treating or diagnosing diseases including metabolic disorders such as
CC diabetes and obesity, infectious diseases, anorexia, cancer,
CC cardiovascular diseases including hypertension and atherosclerosis,
CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC Furthermore, the nucleic acids and polypeptides may also be used to
CC identify molecules that modulate or inhibit neurogenesis, cell
CC differentiation and proliferation, haemopoiesis, wound healing and
CC angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC be used as hybridisation probes, in chromosome mapping, tissue typing,
CC preventive medicine and pharmacogenomics. The current sequence is that of

CC the human NOV cDNA of the invention.

XX
SQ Sequence 838 BP; 208 A; 221 C; 190 G; 219 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.23e-18	Length:	838
Score:	225.50	Matches:	40
Percent Similarity:	60.8%	Conservative:	33
Best Local Similarity:	33.3%	Mismatches:	42
Query Match:	28.2%	Indels:	5
DB:	10	Gaps:	2

US-10-664-025-3903 (1-153) x ADE28662 (1-838)

QY	5	ValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeuHisValLysIle	24
DB	124	CTGATGATCTCTCTCACCTTGGCGCTTTCAGCCGGTTTCAGGACAAAGTCCAATGACTGTG	183
QY	25	SerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeu	44
DB	184	CTGTGCTCCATAGACTGGTTTCATGGTCACAGTCACCCCTTCATGCTAAACAACGATGTG	243
QY	45	TyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThr	64
DB	244	TGTGTACACTTTCATGAACCTACACTTGGCGCTTGGGTTGCCCCCAACCATGTTACGCCA	303
QY	65	TyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSer	84
DB	304	CACGCCTACCACTTACCTACCTACCTGTTTACTGAATGTGGCATCAGGGCCAAAGCTGTCTCT	363
QY	85	GluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAsp	104
DB	364	CAGGACATGGTTATCTACACACTGAGATACACTCTCTCTTAAGGCG-----ACG	414
QY	105	ProGlnGlu-----IleHisLeuGluCysSerThrSerArgLysSerValTrpLeuThr	122
DB	415	CCATCTAAGTTTGTGTATCCAGTGTCATGTGCTGCTGCCCCCAAAAGTCCCATGGCTCACC	474

Search completed: May 9, 2006, 13:30:07
Job time : 801.6 secs

This Page Blank (uspto)

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 13:33:17 ; Search time 200.7 Seconds
(without alignments)
2032.638 Million cell updates/sec

Title: US-10-664-025-3903
Perfect score: 799
Sequence: 1 MALEVLMLLAVLIWTGAENL.....PSFFIADFTTAELGLLIF 153

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn2n.model -DEV=xlp
-Q-/abes/ABSSWEB spool/US10664025/runat_08052006_173649_29466/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abes03p -USER=US10664025@CGN_1_1_359 @runat_08052006_173649_29466
-NCFU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/prodata/1/ina/1 COMB.seq.*
2: /cgn2_6/prodata/1/ina/5 COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/H COMB.seq.*
6: /cgn2_6/prodata/1/ina/ECTUS COMB.seq.*
7: /cgn2_6/prodata/1/ina/PP COMB.seq.*
8: /cgn2_6/prodata/1/ina/RE COMB.seq.*
9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	799	100.0	512	3	US-09-621-976-43
2	789	98.7	642	3	US-09-621-976-1362
3	785	98.2	1492	3	US-10-104-047-1704
4	608	76.1	560	3	US-09-621-976-2587
5	468	58.6	372	3	US-09-621-976-15700
6	429	53.7	469	3	US-09-621-976-1363
7	223.5	28.0	1129	3	US-09-949-016-5361
8	223.5	28.0	96590	3	US-09-949-016-17103
9	142.5	17.8	1098	3	US-09-484-970B-55

10	86.5	10.8	23222	3	US-09-949-016-15949	Sequence 15949, A
11	86.5	10.8	25922	3	US-09-949-016-11874	Sequence 11874, A
12	85.5	10.7	2655	3	US-09-184-418C-95	Sequence 95, Appl
13	85.5	10.7	2655	3	US-10-290-579A-95	Sequence 95, Appl
14	85.5	10.7	2721	3	US-09-184-418C-96	Sequence 96, Appl
15	85.5	10.7	2721	3	US-10-290-579A-96	Sequence 96, Appl
16	85.5	10.7	9060	3	US-09-184-418C-10	Sequence 10, Appl
17	85.5	10.7	9060	3	US-10-290-579A-10	Sequence 10, Appl
18	83	10.4	769	3	US-09-270-767-909	Sequence 909, App
19	83	10.4	769	3	US-09-270-767-16191	Sequence 16191, A
20	83	10.4	29686	3	US-09-949-016-16379	Sequence 16379, A
21	82	10.3	10091	3	US-09-657-013-24	Sequence 24, Appl
22	82	10.3	10182	3	US-09-657-013-25	Sequence 25, Appl
23	82	10.3	10182	3	US-09-657-013-27	Sequence 27, Appl
24	82	10.3	10182	3	US-09-657-013-28	Sequence 28, Appl
25	82	10.3	10182	3	US-09-949-016-4338	Sequence 4338, Ap
26	82	10.3	79858	3	US-09-949-016-16080	Sequence 16080, A
27	81	10.1	15192	3	US-09-949-016-15143	Sequence 15143, A
28	79	9.9	1721	3	US-09-919-039-315	Sequence 315, App
29	78.5	9.8	37838	3	US-09-949-016-12134	Sequence 12134, A
30	78.5	9.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
31	78.5	9.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
32	76.5	9.6	1086	3	US-09-489-039A-1370	Sequence 1370, Ap
33	76.5	9.6	36720	3	US-09-949-016-13476	Sequence 13476, A
34	76	9.5	477	3	US-09-583-110-587	Sequence 587, App
35	76	9.5	498	3	US-09-107-433-1351	Sequence 1351, Ap
36	76	9.5	852	3	US-08-961-527-339	Sequence 339, App
37	76	9.5	972	3	US-09-489-039A-2991	Sequence 2991, Ap
38	75.5	9.4	885	2	US-08-628-291-1	Sequence 1, Appli
39	75.5	9.4	885	2	US-09-128-722-1	Sequence 1, Appli
40	75.5	9.4	885	2	US-09-326-157-2	Sequence 2, Appli
41	75.5	9.4	885	3	US-10-135-755-2	Sequence 2, Appli
42	75.5	9.4	1158	3	US-09-328-352-3761	Sequence 3761, Ap
43	75.5	9.4	1269	3	US-09-949-016-4702	Sequence 4702, Ap
44	75.5	9.4	1280	2	US-08-628-291-15	Sequence 15, Appl
45	75.5	9.4	1280	2	US-09-128-722-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-43
; Sequence 43, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 43
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..511
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.6999998092651
; OTHER INFORMATION: seq VLMLLAVLIWTGA/EN

US-09-621-976-43

Alignment Scores:
Pred. No.: 4.29e-98 Length: 512
Score: 799.00 Matches: 153
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0

```
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-10-664-025-3903 (1-153) x US-09-621-976-43 (1-512)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrrThrGlyAlaGluAsnLeu 20
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 112
QY 21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 113 CATGTGAAAATAAGTTGCTCTCGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 172
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAAATCTGTATATATTTGCGGATGAAATACATCTGGGAATGGGCTGCCCTGCAAA 232
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA 292
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 293 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrr 120
DB 353 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACCAAGTTTCTACTGAGAAATGAATAAATTTGGATCTCTAGTCCITTTATTGCTGAC 472
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeuPhe 153
DB 473 TTTTCAGACACAGCAGAGAGTAGGATTATTAATCTTC 511
RESULT 2
US-09-621-976-1362
; Sequence 1362, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1362
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 53..526
; NAME/KEY: sig_peptide
; LOCATION: 53..103
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.6999998092651
; OTHER INFORMATION: seq VLMMLAVLIWTGA/EN
US-09-621-976-1362
Alignment Scores:
Pred. No.: 1,33e-96 Length: 642
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 3 Gaps: 0
US-10-664-025-3903 (1-153) x US-09-621-976-1362 (1-642)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrrThrGlyAlaGluAsnLeu 20
DB 53 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 112
QY 21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 113 CATGTGAAAATAAGTTGCTCTCGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 172
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 173 AGCAGAAATCTGTATATATTTGCGGATGAAATACATCTGGGAATGGGCTGCCCTGCAAA 232
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 233 CGGATACATACATATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA 292
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 293 AGGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 352
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrr 120
DB 353 ATAGATCATGACCTCAGGAAATCCATTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 412
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 413 CTTACACCAAGTTTCTACTGAGAAATGAATAAATTTGGATCTCTAGTCCITTTATTGCTGAC 472
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeuPhe 151
DB 473 TTTTCAGACACAGCAGAGAGTAGGATTATTA 505
RESULT 3
US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1704
Alignment Scores:
Pred. No.: 1.6e-95 Length: 1492
Score: 785.00 Matches: 150
Percent Similarity: 99.3% Conservativeness: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 98.2% Indels: 0
DB: 3 Gaps: 0
US-10-664-025-3903 (1-153) x US-10-104-047-1704 (1-1492)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrrThrGlyAlaGluAsnLeu 20
DB 52 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGTGATTTGGACCGGTGCTGAGAACCTC 111
QY 21 HisVallyIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 112 CATGTGAAAATAAGTTGCTCTCGACTGGTGTGATGGCCTCAGTTATCCAGTTGCAGAA 171
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
```

Db 172 AGCAGAAATCTGTATATATTTGGGATGATTAACATCTGGATGGCTGCCCTGCANAT 231
QY 61 AtqIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 232 CGGATACATACATATGTATATAGTTTATATATCTTTGTCGTGATTTGGCATCAGGACA 291
QY 81 AtqValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db 292 AGGTAAGTTCTCAGGAAATCTCTCTTTTCAAACGAGCTGTACTTTACCCCAAGGAAT 351
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIlyssSerValTyr 120
Db 352 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTCCACCTCTAGGAAATCAGTGTGG 411
QY 121 LeuThrProValSerThrGluAsnGluIleLeuAspProSerProPheIleAlaAsp 140
Db 412 CTTACACCAAGTTTCTACTGAGATGAATAAAATTTGGATCCTAGTCTCTTTATTGCTGAC 471
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db 472 TTTACAGAACACGACAGAGATTAGGATTATTA 504

RESULT 4

US-09-621-976-2587
; Sequence 2587, Application US/09621976
; Patent No. 6639063

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm

; SEQ ID NO 2587
; LENGTH: 560

; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: CDS

; LOCATION: 372..539
US-09-621-976-2587

Alignment Scores:

Pred. No.: 2,52e-72 Length: 560
Score: 608.00 Matches: 141
Percent Similarity: 88.1% Conservatives: 0
Best Local Similarity: 88.1% Mismatches: 10
Query Match: 76.1% Indels: 9
DB: 3 Gaps: 2

US-10-664-025-3903 (1-153) x US-09-621-976-2587 (1-560)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTyrThrGlyAlaGluAsnLeu 20
Db 53 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGACCGGTGCTGAGAACCTC 112
QY 21 HisValValIleSerCysSerLeuAsp---TyrLeuMetValSerValIleProValala 39
Db 113 CATGTGAANAATAAGTTGCTCTCTGGGACTTGTTGATGGTCTCAGTTATCCAGTTGCA 172
QY 40 GluSerArgAsnLeu-TyrIlePhe---AlaAspGluLeuHisLeuGly-MetGlyCysP 58
Db 173 GAAACAGAAATCTGKAWAAATTTGGSGWTGRAWTTAAMWATCTGGGAAATGGGCTGCC 232
QY 58 roAlaAsnArg-IleHisThrTyrValTyrGlu-PheIleTyrLeuValArgAspCysG 77
Db 233 CTGCAATCGGGATACATACATATGATGAAGTTTATATATCTTGTTCGTGATTGTGG 292
QY 77 yIleArgThrArg-ValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheT 97

Db 293 CATCAGACAGGGYAGTTTCTGAGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTA 352
QY 97 hrPro-ArgAsnIleAspHisAspProGlnGluIleHisLeu-GluCysSerThrSerAr 116
Db 353 CCCCAAAGGAATATAGATCATGACCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAG 412
QY 116 glyssSerValTyrLeuThrProValSerThrGluAsnGluIleLeuAspProSerPr 136
Db 413 GAAATCAGTGTGGCTTTACACCAAGTTTCTACTGAGAAATGAATAAAATTTGGATCCAAGTCC 472
QY 136 oPheIleAlaAspPheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db 473 TTTTATGTGACTTTTCAGAACACGACAGAGATTAGGATTATTA 518

RESULT 5

US-09-621-976-15700
; Sequence 15700, Application US/09621976
; Patent No. 6639063

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm

; SEQ ID NO 15700
; LENGTH: 372

; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: 351
; OTHER INFORMATION: n=a, g, c o r t

US-09-621-976-15700

Alignment Scores:

Pred. No.: 9,7e-54 Length: 372
Score: 468.00 Matches: 97
Percent Similarity: 97.0% Conservatives: 0
Best Local Similarity: 97.0% Mismatches: 2
Query Match: 58.6% Indels: 2
DB: 3 Gaps: 0

US-10-664-025-3903 (1-153) x US-09-621-976-15700 (1-372)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTyrThrGlyAlaGluAsnLeu 20
Db 76 ATGGCGTTAGAAAGTCTTGATGCTCTCTCTCT-GTCTTTGATTTGACCGGTGCTGAGAACCTC 134
QY 21 HisValValIleSerCysSerLeuAspTyrLeuMetValSerValIleProValalaGlu 40
Db 135 CATGTGAANAATAAGTTGCTCTCTGACCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 194
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 195 AGCAGAAATCTGATATATATTTGCGGATGAATACATCTGGGAAATGGGCTGCCCTGCANAT 254
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 255 CGGATACATACATATGATGATGATATATATCTTGTTCGTGATTTGTCATGTCATGCAACA 314
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThr-GluLeuTyrPheThrProArg 99
Db 315 AGGGTAGTTTCTGAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGG 372

RESULT 6

US-09-621-976-1363
; Sequence 1363, Application US/09621976
; Patent No. 6639063

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Robert S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1363
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VLMLLAVLIWGA/EN
US-09-621-976-1363
```

```
Alignment Scores:
Pred. No.: 2,43e-48 Length: 469
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 53.7% Indels: 0
DB: 3 Gaps: 0
```

US-10-664-025-3903 (1-153) x US-09-621-976-1363 (1-469)

```
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeu 20
Db 137 ATGGCGTTAGAACTCTGATGCTCCTCGCTGCTTGAATGGACCGGTGCTGAGAACCTC 196
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 197 CATGGAAATAAGTTGCTCTCGACTGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA 256
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 257 AGCAGAAATCTGATATATTTCGGATGAATACATCTGGGAATGGCTGCCCTGCAAT 316
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 317 CGGATACATACATATGATGATATATATATATATATATATATATATATATATATATAT 376
QY 81 ArgVal 82
Db 377 AGGGTA 382
```

```
RESULT 7
US-09-949-016-5361
; Sequence 5361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 5361
; LENGTH: 1129
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5361
```

```
Alignment Scores:
Pred. No.: 4.11e-20 Length: 1129
Score: 223.50 Matches: 40
Percent Similarity: 60.0% Conservative: 32
Best Local Similarity: 33.3% Mismatches: 43
Query Match: 28.0% Indels: 5
DB: 3 Gaps: 2
```

US-10-664-025-3903 (1-153) x US-09-949-016-5361 (1-1129)

```
QY 5 ValLeuMetLeuLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeuHisValIle 24
Db 314 CTGATGATCTCTCCTCACCTCTCGGTTTTTCAGCGGTTTCAGGACAAAGTCCAAATGACTGTG 373
QY 25 SerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeu 44
Db 374 CTGTCTCCATAGACTGTTCATGTGCACGTGCACCCCTTCATCTCTAAACACCATGTG 433
QY 45 TyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThr 64
Db 434 TGTGTACACTTTCATGAACACTACACTTGGGCTGGGTTGCCCCCAACCATGTCAGCCA 493
QY 65 TyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSer 84
Db 494 CACGCTACCAAGTTCCACTCCGTTTACTGAATGTGCATCAGGCGCCCAAGCTGTCTCT 553
QY 85 GluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAsp 104
Db 554 CAGGACATGGTTATCTACAGCACTGAGATACACTACTCTTTAAGGGC-----ACG 604
QY 105 ProGlnGlu-----IleHisLeuGluCysSerThrSerArgLysSerValTrpLeuThr 122
Db 605 CCATCTAAGTTTGTGATCCAGTGTGTCGCCCCCAAAAGTCCCCCATGGCTCACC 664
```

```
RESULT 8
US-09-949-016-17103
; Sequence 17103, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 17103

LENGTH: 96690

TYPE: DNA

ORGANISM: Human

US-09-949-016-17103

Alignment Scores:

Pred. No.: 2.94e-17 Length: 96690

Score: 223.50 Matches: 40

Percent Similarity: 60.0% Conservative: 32

Best Local Similarity: 33.3% Mismatches: 43

Query Match: 28.0% Indels: 5

DB: 3 Gaps: 2

; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 95
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94CY017.41; gene=tat
US-10-290-579A-95

Alignment Scores:
Pred. No.: 0.549 Length: 2655
Score: 85.50 Matches: 37
Percent Similarity: 41.2% Conservative: 26
Best Local Similarity: 24.2% Mismatches: 44
Query Match: 10.7% Indels: 46
DB: 3 Gaps: 7

US-10-664-025-3903 (1-153) x US-10-290-579A-95 (1-2655)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrrThr-----GlyAlaGlu 18
Db 269 GTAGGACTGATAGCTTTAATCTTAGCAATAGTTGTATGGACTATAGTATTCATAGAA 328
QY 19 AsnLeuHisValIysIleSerCysSerLeuAspTrrLeuMetValSerValIleProVal 38
Db 329 TATAAGAAAATTAAAGAACAAAGGAAAATAGACTGGTTAATCAAAAGAAATAGTGAGAGA 388
QY 39 AlaGlu-----SerArgAsnLeuTyrIlePheAlaAsp 49
Db 389 GCAGAGACAGTCGGCAATGAGAGTGATGGGCACACAGAGAACTATCAGCACTTGTGGAG 448
QY 50 GluLeuHisLeuGly-MetGly-----CysProAlaAsnArgIleHisThrTy 65
Db 449 AGGGGCACTCTTGATTTTGGGGATGTTAATATGTGTAAGCTACAGATTTGTGGGTGCAC 508
QY 65 rValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSerGI 85
Db 509 AGTACTATGAGTACCTGTGTGGAAGATGCA-----GlnThrGluLeuTyrPheTh 97
QY 85 uGluThrLeuLeuPhe-----GlnThrGluLeuTyrPheTh 97
Db 543 -GATACCATCTCTATTGTTGTCATCAGATGCTAAAGCATATGATACAGAAAGTGCAT----- 596
QY 97 rProArgAsnIle-----AspHisAspProGlnGluI 108
Db 597 -----AATGTATGGGCCACACATGCTGTGTACCCACAGACCCCAACCAAGAAAT 649
QY 108 eHisLeuGluCysSerThrSerArgIysSerValTrr 120
Db 650 AAACCTGGAAAATGTAACAGAAAATTTTAAATATGTGG 686

RESULT 14
US-09-184-418C-96
; Sequence 96, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 96

; LENGTH: 2721
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94CY017.41; gene=rev
US-09-184-418C-96

Alignment Scores:
Pred. No.: 0.57 Length: 2721
Score: 85.50 Matches: 37
Percent Similarity: 41.2% Conservative: 26
Best Local Similarity: 24.2% Mismatches: 44
Query Match: 10.7% Indels: 46
DB: 3 Gaps: 7

US-10-664-025-3903 (1-153) x US-09-184-418C-96 (1-2721)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrrThr-----GlyAlaGlu 18
Db 130 GTAGGACTGATAGCTTTAATCTTAGCAATAGTTGTATGGACTATAGTATTCATAGAA 189
QY 19 AsnLeuHisValIysIleSerCysSerLeuAspTrrLeuMetValSerValIleProVal 38
Db 190 TATAAGAAAATTAAAGAACAAAGGAAAATAGACTGGTTAATCAAAAGAAATAGTGAGAGA 249
QY 39 AlaGlu-----SerArgAsnLeuTyrIlePheAlaAsp 49
Db 250 GCAGAGACAGTCGGCAATGAGAGTGATGGGCACACAGAGAACTATCAGCACTTGTGGAG 309
QY 50 GluLeuHisLeuGly-MetGly-----CysProAlaAsnArgIleHisThrTy 65
Db 310 AGGGGCACTCTTGATTTTGGGGATGTTAATATGTGTAAGCTACAGATTTGTGGGTGCAC 369
QY 65 rValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSerGI 85
Db 370 AGTACTATGAGTACCTGTGTGGAAGATGCA-----GlnThrGluLeuTyrPheTh 97
QY 85 uGluThrLeuLeuPhe-----GlnThrGluLeuTyrPheTh 97
Db 404 -GATACCATCTCTATTGTTGTCATCAGATGCTAAAGCATATGATACAGAAAGTGCAT----- 457
QY 97 rProArgAsnIle-----AspHisAspProGlnGluI 108
Db 458 -----AATGTATGGGCCACACATGCTGTGTACCCACAGACCCCAACCAAGAAAT 510
QY 108 eHisLeuGluCysSerThrSerArgIysSerValTrr 120
Db 511 AAACCTGGAAAATGTAACAGAAAATTTTAAATATGTGG 547

RESULT 15
US-10-290-579A-96
; Sequence 96, Application US/10290579A
; Patent No. 6897301
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; FILE REFERENCE: D6287D
; CURRENT APPLICATION NUMBER: US/10/290,579A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 09/184,418
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 270
; SEQ ID NO 96
; LENGTH: 2721
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94CY017.41; gene=rev
US-10-290-579A-96

Alignment Scores:

Pred. No.:	0.57	Length:	2721
Score:	85.50	Matches:	37
Percent Similarity:	41.2%	Conservative:	26
Best Local Similarity:	24.2%	Mismatches:	44
Query Match:	10.7%	Indels:	46
DB:	3	Gaps:	7

US-10-664-025-3903 (1-153) x US-10-290-579A-96 (1-2721)

```
Qy 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTrrpThr-----GlyAlaGlu 18
   :: ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 130 GTAGGACTGATGACTGCTTAACTTAGCAATAGTTGTATGGACTATAGTATTTCATAGAA 189

Qy 19 AsnLeuHisValLysIleSerCysSerLeuAspTrrpLeuMetValSerValIleProVal 38
   ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 190 TATAAGAAATTTAAGAACAAAGGAAATAGACTGGTTAATCAAAAGAAATAAGTGAGAGA 249

Qy 39 AlaGlu-----SerArgAsnLeuTyrrilePheAlaAsp 49
   ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 250 GCAGAGACAGTGGCAATCAGAGTGTATGGGGACACAGAGGAACATATCAGCACTTGTGGAG 309

Qy 50 GluLeuHisLeuGly-MetGly-----CysProAlaAsnArgIleHisThrTy 65
   ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 310 AGGGGGCATCTTGATTTTGGGGATGTTAATAATGTGTAAAGCTACAGATTTTGTGGTCCAC 369

Qy 65 rValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThrArgValValSerGl 85
   ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 370 AGTATACTATGGAGTACCTGTGTGGAAGATGCA----- 403

Qy 85 uGluThrLeuLeuPhe-----GlnThrGluLeuTyrrPheTh 97
   ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 404 -GATACCATCCTATTTTTGTGCATCAGATGCTAAAGCATATGATACAGAGTGCAT----- 457

Qy 97 rProArgAsnIle-----AspHisAspProGlnGluIl 108
   ||| ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 458 -----AATGTATGGGCCACACATGCCTGTGTACCACAGACCCCAACCCACAAGAAAT 510

Qy 108 eHisLeuGluCysSerThrSerArgLysSerValTrp 120
   ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 511 AAACCTGGAAATGTAACACAGAAAAATTTTAAATATGTGG 547
```

Search completed: May 9, 2006, 14:54:01
Job time : 228.7 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: May 9, 2006, 14:40:21 ; Search time 1012.5 Seconds
(without alignments)
1874.392 Million cell updates/sec

Title: US-10-664-025-3903
Perfect score: 799
Sequence: 1 MALEVLMLAVLIWTGAENL.....PSPFIADFTQAEELGLLIP 153

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DBV=xl
-Q=/abss/ABSSWEB.spool/US10664025/runat.08052006.173653.29528/app_query.fasta_1
-DB=Published Applications NA Main -OFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs06p
-USER=US10664025 @CGN 1.1 2064 @runat.08052006.173653.29528 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	789	98.7	691 3 US-09-731-872-19 Sequence 19, Appl
2	789	98.7	691 3 US-09-731-872-19 Sequence 19, Appl
3	789	98.7	691 3 US-09-731-872-19 Sequence 19, Appl
4	789	98.7	1337 8 US-10-643-836-24 Sequence 2, Appl
5	789	98.7	1480 7 US-10-467-046-2 Sequence 16, Appl
6	789	98.7	1898 5 US-10-467-535-16 Sequence 27, Appl
7	789	98.7	1898 7 US-10-050-704-27 Sequence 27, Appl

8	789	98.7	1898	9	US-10-472-533-175	Sequence 175, App
9	785	98.2	1492	6	US-10-104-047-1704	Sequence 1704, Ap
10	429	53.7	470	3	US-09-731-872-24	Sequence 24, Appl
11	429	53.7	470	3	US-09-876-997-24	Sequence 24, Appl
12	429	53.7	470	3	US-10-643-836-24	Sequence 24, Appl
13	325.5	40.7	5013	3	US-09-984-429-532	Sequence 532, App
14	325	40.7	10115	8	US-10-467-046-1	Sequence 1, Appli
15	225.5	28.2	838	7	US-10-287-971-39	Sequence 39, Appl
16	224.5	28.1	589	8	US-10-333-872A-44	Sequence 44, Appl
17	223.5	28.0	1148	6	US-10-411-224-36	Sequence 36, Appl
18	223.5	28.0	1148	6	US-10-047-021-36	Sequence 36, Appl
19	223.5	28.0	1148	9	US-10-970-493-36	Sequence 36, Appl
20	223.5	28.0	1153	3	US-09-745-763-202	Sequence 202, App
21	223.5	28.0	1187	3	US-09-996-952-4	Sequence 4, Appli
22	167	20.9	386	5	US-10-040-739-196	Sequence 196, App
23	161	20.2	697	7	US-10-287-971-41	Sequence 41, Appl
24	98.5	12.3	4490	9	US-10-450-763-12039	Sequence 12039, A
C 25	95	11.9	472	3	US-09-864-761-3678	Sequence 1678, Ap
C 26	88.5	11.1	413	3	US-09-864-761-5976	Sequence 5976, Ap
C 27	88.5	11.1	705	9	US-10-450-763-27333	Sequence 27333, A
C 28	88	11.0	106378	5	US-10-087-192-1624	Sequence 1624, Ap
C 29	88	11.0	368003	8	US-10-719-993-6805	Sequence 6805, Ap
30	85.5	10.7	743	3	US-09-925-297-143	Sequence 143, App
31	85.5	10.7	2655	6	US-10-290-579-95	Sequence 95, Appl
32	85.5	10.7	2855	10	US-11-135-597-95	Sequence 95, Appl
33	85.5	10.7	2721	6	US-10-290-579-96	Sequence 96, Appl
34	85.5	10.7	2721	10	US-11-135-597-96	Sequence 96, Appl
35	85.5	10.7	9060	6	US-10-290-579-10	Sequence 10, Appl
36	85.5	10.7	9060	10	US-11-135-597-10	Sequence 10, Appl
37	84	10.5	3723	7	US-10-282-122A-9832	Sequence 9832, Ap
38	83	10.4	574	4	US-09-925-065A-467415	Sequence 467415,
C 39	82	10.3	10182	9	US-10-504-173-111	Sequence 111, App
C 40	82	10.3	10470	3	US-09-814-353-20168	Sequence 20168, A
41	82	10.3	181343	8	US-10-723-860-2392	Sequence 2392, Ap
42	82	10.3	181343	9	US-10-756-149-2215	Sequence 2215, Ap
C 43	80	10.0	32816	3	US-09-729-094-3	Sequence 3, Appli
C 44	80	10.0	32816	6	US-10-435-631-3	Sequence 3, Appli
C 45	79.5	9.9	620	4	US-09-925-065A-691324	Sequence 691324,

ALIGNMENTS

RESULT 1
US-09-731-872-19
; Sequence 19, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueret, Lydie
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; NAME/KEY: sig_peptide
; LOCATION: 42..515
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLAVLIWTGA/EN
US-09-731-872-19

```

; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLLAVLIWTGA/EN
US-09-876-997-19

Alignment Scores:
Pred. No.: 3,48e-105 Length: 691
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: Gaps: 0

US-10-664-025-3903 (1-153) x US-09-876-997-19 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuLeuTrpThrGlyAlaGluAsnLeu 20
   |||
Db 42 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGGATTTGGACCGGTGCTGAGAACCTC 101
   |||
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
   |||
Db 102 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 161
   |||
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
   |||
Db 162 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 221
   |||
QY 61 ArgIleHisThrTyrlleValTyrllePheIleValLeuValArgAspCysGlyIleArgThr 80
   |||
Db 222 CGGATACATACATATATATATGATGATTTATATATCTTGTCTGATTTGGCATCAGGACA 281
   |||
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
   |||
Db 282 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 341
   |||
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
   |||
Db 342 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 401
   |||
QY 121 LeuThrProValSerThrGluAsnGluLysLeuAspProSerProPheIleAlaAsp 140
   |||
Db 402 CTTACACCAAGTTTCTACTGAGAATGAAATAAAATTGGATCTTGTCTTTATTGCTGAC 461
   |||
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
   |||
Db 462 TTTTCAGACACACAGCAGAGAGTTAGGATTATTA 494
   |||

RESULT 3
US-10-643-836-19
; Sequence 19, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; LOCATION: 42..515

; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLLAVLIWTGA/EN
US-09-876-997-19

Alignment Scores:
Pred. No.: 3,48e-105 Length: 691
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: Gaps: 0

US-10-664-025-3903 (1-153) x US-09-731-872-19 (1-691)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuLeuTrpThrGlyAlaGluAsnLeu 20
   |||
Db 42 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTTGGATTTGGACCGGTGCTGAGAACCTC 101
   |||
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
   |||
Db 102 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 161
   |||
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
   |||
Db 162 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 221
   |||
QY 61 ArgIleHisThrTyrlleValTyrllePheIleValLeuValArgAspCysGlyIleArgThr 80
   |||
Db 222 CGGATACATACATATATATGATGATTTATATATCTTGTCTGATTTGGCATCAGGACA 281
   |||
QY 81 ArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
   |||
Db 282 AGGGTAGTTTCTGAGGAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAAT 341
   |||
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
   |||
Db 342 ATAGATCATGACCTCAGGAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 401
   |||
QY 121 LeuThrProValSerThrGluAsnGluLysLeuAspProSerProPheIleAlaAsp 140
   |||
Db 402 CTTACACCAAGTTTCTACTGAGAATGAAATAAAATTGGATCTTGTCTTTATTGCTGAC 461
   |||
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
   |||
Db 462 TTTTCAGACACACAGCAGAGAGTTAGGATTATTA 494
   |||

RESULT 2
US-09-876-997-19
; Sequence 19, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..515
; NAME/KEY: sig_peptide
; LOCATION: 42..92
; OTHER INFORMATION: Von Heijne matrix
; LOCATION: 42..515
```

```
;
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 42-92
; OTHER INFORMATION: Von Heine matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VLMLAVLIWTGA/EN
US-10-643-836-19

Alignment Scores:
Pred. No.: 3.48e-105 Length: 691
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-643-836-19 (1-691)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleThrThrGlyAlaGluAsnLeu 20
Db 42 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCTGAGAACCTC 101
Qy 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 102 CATGTGAAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 161
Qy 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 162 AGCAGAAATCTGTATATATTTGGGATGAATTAATCATCTGGGAATGGGCTGCCCTGCAAA 221
Qy 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 222 CGGATACATACATATGATATAGTTTATATATCTTGTTCGATTTGGCATCAGGACA 281
Qy 81 ArgValValSerGluGluThrLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db 282 AGGAGTAGTTTCTGAGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGAA 341
Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Db 342 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACACCTCTAGGAAATCAGTGTG 401
Qy 121 LeuThrProValSerThrGluAsnGluLeuLysLeuAspProSerProPheIleAlaAsp 140
Db 402 CTTACACCAAGTTTCTACTGGAATGAATAAATTTGATCTCTAGTCTCTTTATTGCTGAC 461
Qy 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db 462 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 494
```

```
RESULT 4
US-10-467-046-2
; Sequence 2, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Cid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-Potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSPSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53) .. (103)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (104) .. (526)
US-10-467-046-2

Alignment Scores:
Pred. No.: 9.68e-105 Length: 1337
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-467-046-2 (1-1337)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleThrThrGlyAlaGluAsnLeu 20
Db 53 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCTGAGAACCTC 112
Qy 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 113 CATGTGAAAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 172
Qy 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 173 AGCAGAAATCTGTATATATTTGGGATGAATTAATCATCTGGGAATGGGCTGCCCTGCAAA 232
Qy 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 233 CGGATACATACATATGATATAGTTTATATATCTTGTTCGATTTGGCATCAGGACA 292
Qy 81 ArgValValSerGluGluThrLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db 293 AGGAGTAGTTTCTGAGAAACTCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGAA 352
Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Db 353 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACACCTCTAGGAAATCAGTGTG 412
Qy 121 LeuThrProValSerThrGluAsnGluLeuLysLeuAspProSerProPheIleAlaAsp 140
Db 413 CTTACACCAAGTTTCTACTGGAATGAATAAATTTGATCTCTAGTCTCTTTATTGCTGAC 472
Qy 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db 473 TTTTCAGACACAGCAGAGAGTTAGGATTATTA 505
```

```
RESULT 5
US-10-467-535-16
; Sequence 16, Application US/10467535
; Publication No. US20040146970A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; YAO, Monique G.
; APPLICANT: ISON, Craig H.; LU, Yan
; APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
; APPLICANT: BAUGHN, Mariah R.; DING, Li
; APPLICANT: XU, Yuming; GIETZEN, Kimberly J.
; APPLICANT: TANG, Tom Y.; LAL, Preeti G.
; APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
; APPLICANT: LU, Dyung Aina M.; RICHARDSON, Thomas W.
; APPLICANT: TRAM, Oyen K.; KHARE, Reena
; APPLICANT: CHAWLA, Narinder K.
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-0903 USN
; CURRENT APPLICATION NUMBER: US/10/467,535
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/US02/03715
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/268,111
; PRIOR FILING DATE: 2001-02-09
```

; PRIOR APPLICATION NUMBER: US 60/271,175
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,503
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/274,552
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 1480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5284076CB1
US-10-467-535-16

Alignment Scores:
Pred. No.: 1.13e-104 Length: 1480
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-467-535-16 (1-1480)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuLeuThrGlyValAlaGluAsnLeu 20
DB 29 ATGGCGTGTAGAAAGCTTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCTGAGAACCTC 88
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 89 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 148
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 149 AGCAGAAATCTGTATATATTTGGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 208
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 209 CGGATACATACATATGATGAGTTTATATATCTTTGTCGATTTGGCATCAGGACA 268
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 269 AGGTAGTTTCTGAGGAATCTCTCTTTTCAAACCGAGCTGACTTTACCCCAAGGAAT 328
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
DB 329 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 388
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 389 CTTACACCAAGTTCTACTCAGAAATGAAATAAAATTTGGATCTTAGTCTTTATTGCTGAC 448
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
DB 449 TTTCAGACACAGCAGAGAGTTAGGATTATTA 481

RESULT 6

US-10-050-704-27
; Sequence 27, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-050-704-27

Alignment Scores:
Pred. No.: 1.67e-104 Length: 1898
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-050-704-27 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuLeuThrGlyValAlaGluAsnLeu 20
DB 15 ATGGCGTGTAGAAAGCTTTGATGCTCTCGCTGCTTGAATTTGGACCGGTGCTGAGAACCTC 74
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
DB 75 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 134
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
DB 135 AGCAGAAATCTGTATATATTTGGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 194
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
DB 195 CGGATACATACATATGATGAGTTTATATATCTTTGTCGATTTGGCATCAGGACA 254
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
DB 255 AGGTAGTTTCTGAGGAATCTCTCTTTTCAAACCGAGCTGACTTTACCCCAAGGAAT 314
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
DB 315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 374
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
DB 375 CTTACACCAAGTTCTACTCAGAAATGAAATAAAATTTGGATCTTAGTCTTTATTGCTGAC 434
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
DB 435 TTTCAGACACAGCAGAGAGTTAGGATTATTA 467

RESULT 7

US-10-798-512-27
; Sequence 27, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10

; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1398)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-798-512-27

Alignment Scores:
Pred. No.: 1.67e-104 Length: 1898
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 7 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-798-512-27 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
|||
Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCTGAGAACCTC 74
|||
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
|||
Db 75 CATGTGNAATAAGTTGCTCTCGGACTGGTGTGATGCTCAGTTATCCAGTTGCAGAA 134
|||
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
|||
Db 135 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTGGGAATGGCTGCCCTGCAAAAT 194
|||
QY 61 ArgIleHisThrTrpValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80
|||
Db 195 CGGATACATACATATGATATGATGTTTATATATCTTGTTCGTGATTGTCGATCAGGACA 254
|||
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
|||
Db 255 AGGGTAGTTTCTGAGGAACTCTCTTTTCAACCCGAGCTGTACTTTACCCCAAGGAAT 314
|||
QY 101 IleAspHisAspProGlnGluIleHisLeuGlyCysSerThrSerArgLysSerValTrp 120
|||
Db 315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 374
|||
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
|||
Db 375 CTTACACCAAGTTTCTACTGAGATGAATAAAATTTGGATCTAGTCTCTTTATTGCTGAC 434
|||
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
|||
Db 435 TTTCAGACAACACAGAGAGATTAGGATTATTA 467

RESULT 8

US-10-472-533-175
; Sequence 175, Application US/10472533
; Publication No. US20050197285A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS906PCT
; CURRENT APPLICATION NUMBER: US/10/472,533

; CURRENT FILING DATE: 2003-09-20
; PRIOR APPLICATION NUMBER: US 60/331,287
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/306,171
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/277,340
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 650
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1398)..(1398)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1428)..(1428)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-472-533-175

Alignment Scores:
Pred. No.: 1.67e-104 Length: 1898
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 9 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-472-533-175 (1-1898)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
|||
Db 15 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTCTGATTTGGACCGGTGCTGAGAACCTC 74
|||
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
|||
Db 75 CATGTGNAATAAGTTGCTCTCGGACTGGTGTGATGCTCAGTTATCCAGTTGCAGAA 134
|||
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
|||
Db 135 AGCAGAAATCTGTATATATTTCGGATGAATTACATCTGGGAATGGCTGCCCTGCAAAAT 194
|||
QY 61 ArgIleHisThrTrpValTyrlleGluPheIleTyrlleValArgAspCysGlyIleArgThr 80
|||
Db 195 CGGATACATACATATGATATGATGTTTATATATCTTGTTCGTGATTGTCGATCAGGACA 254
|||
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrllePheThrProArgAsn 100
|||
Db 255 AGGGTAGTTTCTGAGGAACTCTCTTTTCAACCCGAGCTGTACTTTACCCCAAGGAAT 314
|||
QY 101 IleAspHisAspProGlnGluIleHisLeuGlyCysSerThrSerArgLysSerValTrp 120
|||
Db 315 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 374
|||
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
|||
Db 375 CTTACACCAAGTTTCTACTGAGATGAATAAAATTTGGATCTAGTCTCTTTATTGCTGAC 434
|||
QY 141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
|||
Db 435 TTTCAGACAACACAGAGAGATTAGGATTATTA 467

RESULT 9

US-10-104-047-1704
; Sequence 1704, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cdna
; FILE REFERENCE: H1-A0105

; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; score 10.7019149919754
; OTHER INFORMATION: seq VMLLAVLWTGA/EN
US-09-731-872-24
Alignment Scores:
Pred. No.: 1.72e-52 Length: 470
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 53.7% Indels: 0
DB: Gaps: 0
US-10-664-025-3903 (1-153) x US-09-731-872-24 (1-470)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 ATGGCGCTTAAAGATCTTTGATGTCTTCCTCGTGCTTTGATTGGACCCGGTCTGAGAACCCTC 196
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 CATGTGAAATAAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 256
QY 41 SerArgAsnLeuTyrrilePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 AGCAGAAATCTGTATATATTTTGGGATGAATTACATCTGGGAATGGCTGCCCTGCAAAAT 316
QY 61 ArgIleHisThrTrpValTyrrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 CGGATACATACATATGTATATGAGTTTATATATCTTTGTTCTGTAATTTGGCATCAGGACA 376
QY 81 ArgVal 82
| | | | |
Db 377 AGGGTA 382
RESULT 11
US-09-876-997-24
; Sequence 24, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix

```
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-09-876-997-24

Alignment Scores:
Pred. No.: 1.72e-52 Length: 470
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 53.7% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903 (1-153) x US-09-876-997-24 (1-470)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
Db 137 ATGCGGTTAGAAAGTCCTTGGATGCTCTCGCTGCTTGGATTTGGACCGGTGCTGGAACCTC 196

QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 197 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA 256

QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 257 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCCAAAT 316

QY 61 ArgIleHisThrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 317 CGATACATACATATGTATATATATATATATATATATATATATATATATATATATATATAT 376

QY 81 ArgVal 82
Db 377 AGGGTA 382

RESULT 12
US-10-643-836-24
; Sequence 24, Application US/10643836
; Publication No. US20050096458A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 24
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..454
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 137..187
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.7019149919754
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN
US-10-643-836-24

Alignment Scores:
Pred. No.: 1.72e-52 Length: 470
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
```

```
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 53.7% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-643-836-24 (1-470)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
Db 137 ATGCGGTTAGAAAGTCCTTGGATGCTCTCGCTGCTTGGATTTGGACCGGTGCTGGAACCTC 196

QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 197 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGTCTCAGTTATCCCAAGTTGCAGAA 256

QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 257 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCCAAAT 316

QY 61 ArgIleHisThrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 317 CGATACATACATATGTATATATATATATATATATATATATATATATATATATATATATAT 376

QY 81 ArgVal 82
Db 377 AGGGTA 382

RESULT 13
US-09-984-429-532
; Sequence 532, Application US/09984429
; Publication No. US20040010132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P2
; CURRENT APPLICATION NUMBER: US/09/984,429
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,591
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 727
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 532
; LENGTH: 5013
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-984-429-532

Alignment Scores:
Pred. No.: 1.13e-35 Length: 5013
Score: 325.50 Matches: 95
Percent Similarity: 21.8% Conservative: 0
Best Local Similarity: 21.8% Mismatches: 0
Query Match: 40.7% Indels: 341
DB: 3 Gaps: 1

US-10-664-025-3903 (1-153) x US-09-984-429-532 (1-5013)

QY 22 ValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSer 41
```

Db 1 GTGAAATAAGTTGCTCTGACCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAAAGC 60
QY |||||
Db 42 ArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArg 61
QY |||||
Db 61 AGAAATCTGTATATATTTGCGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAATCGG 120
QY |||||
Db 62 IleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArg 81
QY |||||
Db 121 ATACATACATATGATATGAGTTTATATATCTTTGTTGCGTATGTTGGCATGAGGACNAGG 180
QY |||||
Db 81 ----- 81
Db 181 GTAAGAACAGTGATTGTTGTAAAAAATACTGCATGTTTGTGCAGACTTTTATGCTAGT 240
QY ----- 81
Db 241 ATTAATAATGCTTTTAAAAAATAATAGTTTGTCTCTCANTGTAGTAATTTATGGGCAGG 300
QY ----- 81
Db 301 GGACACGGAATTATCATTTATTTTCTACTGAATATGGGTCAATCTGTGTGCACAGGAC 360
QY ----- 81
Db 361 CTTAGATCACAGATTATAAATTTGGCTAGGGAACGATTCTAAGTTGCCCTTAGGCTACT 420
QY ----- 81
Db 421 ATTATATTTTACTAAATATTAGCAGCTGCTTTAAATTTTGTGTGATCTTCTTAATGTGG 480
QY ----- 81
Db 481 AAGTCCACCAATCTTGGGAAGCTCCATCATGTTCAATGTTAGTCATTTGGAATCTATCC 540
QY ----- 81
Db 541 ATGTTCCATGAGGATTTGTGTAAATATATAGTGGATTTTCATTGAAATGAAATCTCTTGT 600
QY ----- 81
Db 601 CATTGAGTAGGCACAAATCTTAAATTTGCTTTATCTGGGATATATCTTCTCACAG 660
QY ----- 81
Db 661 ATTCCCAATTTAAGCAATTTATATTCACACTACACAAATATTTGGGCCAGTGACCAAT 720
QY ----- 81
Db 721 GATAATAACAATCAATTTTGTGAATAAAGTATTTTATTTAGAAATTTGGGAATTTGGAGCTT 780
QY ----- 81
Db 781 CAGAAGGTGCTTTACAGCTTTTCCAGGATGCAGCTGTTGTTGGTGGTAGTTTAGATAC 840
QY ----- 81
Db 841 CAACACGAGTGAATACTCTCCGAAGCTTGTCTTTAAACCACTCTCCGTTATTTGACTT 900
QY ----- 81
Db 901 ACAAAATCACCTTTCTCAAAGGAACGTTGTGGAACCTTTTCATTAATGAATCATAAAGATG 960
QY ----- 81
Db 961 GCTGTAAGCGGAGTTTAAATTTGCTGCAATGAATTAAGATTTGTTTGGCCCCCAGGCATT 1020
QY ----- 81
Db 1021 TTAATTTCACTCTTAGGATAACACTCTCTTCTGTCCAGTGAGTTGTTCATATCTTTAT 1080
QY ----- 81

Db 1081 TTATATTTAAGAATGACAGTAGCCATTTTGGAGATCATGAATGTTTTTACATCTATTTC 1140
QY 81 ----- 81
Db 1141 AAGTGTGTGTTATCTCTCATCTGTGTACTCCAATATATCTTCCACTGAACTTCATGCTCTCC 1200
QY 82 ---ValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
QY |||||
Db 1201 TAGGTAGTTTCTGAGGAAACTCTCCTTTTCAAACCGAGCTGTACTTTTACCCCAAGGAAT 1260
QY |||||
Db 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArg 116
QY |||||
Db 1261 ATAGATCATGACCTCCCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGG 1308
QY |||||
RESULT 14
US-10-467-046-1
; Sequence 1, Application US/10467046
; Publication No. US20040235709A1
; GENERAL INFORMATION:
; APPLICANT: Salter-Gid, Luisa
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Chicca, Barbara A.
; APPLICANT: Chicca, John
; APPLICANT: Yen-potin, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: GSP3 Polynucleotides and Polypeptides and Uses Thereof
; FILE REFERENCE: G-102US03REG
; CURRENT APPLICATION NUMBER: US/10/467,046
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,156
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 10115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2387...2501
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5443...5646
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6649...6747
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8907...9774
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2986)..(2986)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3847)..(3848)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3879)..(3879)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4105)..(4105)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4107)..(4109)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7347)..(7348)
; OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc feature
LOCATION: (8197)..(8197)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (8293)..(8293)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (8334)..(8336)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (8342)..(8349)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (9942)..(9942)
OTHER INFORMATION: n is a, c, g, or t
US-10-467-046-1

Alignment Scores:
Pred. No.: 3.98e-35 Length: 10115
Score: 325.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 40.7% Indels: 0
DB: 8 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-467-046-1 (1-10115)

QY 22 ValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSer 41
Db 5444 GTGAATAAGTTGCTCTCTGGACTGGTTCATGCTCTCAGTTATCCAGTTGCAGAAAGC 5503
QY 42 ArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArg 61
Db 5504 AGAANTCTGATATATTTTGGGATGAATATACATCTGGGAATGGGCTGCCCTGCAAAATCG 5563
QY 62 IleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArg 81
Db 5564 ATACATACATATGATATGATTTATATATCTTTGTCGTGATTTGGCATCAGACAAG 5623
QY 82 val 82
Db 5624 GTA 5626

RESULT 15
US-10-287-971-39
Sequence 39, Application US/10287971
Publication No. US20040067882A1
GENERAL INFORMATION:
APPLICANT: Alsobrook, et al
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-480A
CURRENT APPLICATION NUMBER: US/10/287,971
CURRENT FILING DATE: 2002-11-05
PRIOR APPLICATION NUMBER: 09/997,425
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 10/035,568
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/338,626
PRIOR FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: 60/401,479
PRIOR FILING DATE: 2002-08-06
PRIOR APPLICATION NUMBER: 60/333,072
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: 60/348,283
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: 60/393,262
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 60/406,181

PRIOR FILING DATE: 2002-08-26
NUMBER OF SEQ ID NOS: 397
SOFTWARE: Curaseqlist version 0.1
SEQ ID NO 39
LENGTH: 838
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (100)..(735)
US-10-287-971-39

Alignment Scores:
Pred. No.: 3.61e-22 Length: 838
Score: 225.50 Matches: 40
Percent Similarity: 60.8% Conservative: 33
Best Local Similarity: 33.3% Mismatches: 42
Query Match: 28.2% Indels: 5
DB: 7 Gaps: 2

US-10-664-025-3903 (1-153) x US-10-287-971-39 (1-838)

QY 5 ValLeuMetLeuLeuAlaValLeuIleTyrThrGlyAlaGluAsnLeuHisValLysIle 24
Db 124 CTGATGATCTCTCTCACCTCTGCGCTTCAGCGGTTCAGGACAAAGTCCAATGACTGTG 183
QY 25 SerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeu 44
Db 184 CTGTGCTCCATAGACTGGTTTCATGTCACAGTGCACCCCTTCATGCTAAACACGATGTG 243
QY 45 TyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThr 64
Db 244 TGTGTACACTTTTCATGACACTTGGGCTGGGTGGCCCCCAACCATGTTTCAGCCA 303
QY 65 TyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSer 84
Db 304 CAGGCTACAGTTTCACCTACCGTGTACTGAATGTGGCATCAGGCGCAAAAGTGTCTCT 363
QY 85 GluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAsp 104
Db 364 CAGGACATGGTTATCTACAGCACTGAGATACACTACTCTTCTAAGGC-----ACG 414
QY 105 ProGlnGlu-----IleHisLeuGluCysSerThrSerArgLysSerValTrpLeuThr 122
Db 415 CCATCTAAGTTTGTGATCCAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474

Search completed: May 9, 2006, 15:21:53
Job time : 1017.5 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 14:45:03 ; Search time 682.2 Seconds
(without alignments)
1370.049 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLITGAENL.....PSPFIADFTTAEELGLIP 153

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9306428 segs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10664025/runat_08052006_173656_29587/app_query.fasta.1
-DB=Published Applications NA New -QWMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs808
-USER=US10664025 @CNC 1.1.981 @runat_08052006_173656_29587 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /SID55/ptodata/1/pubpna/US08_NEW_PUB.seq1.*
2: /SID55/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /SID55/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /SID55/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /SID55/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /SID55/ptodata/1/pubpna/US09_NEW_PUB.seq.*
7: /SID55/ptodata/1/pubpna/US09_NEW_PUB.seq1.*
8: /SID55/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
9: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq.*
10: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq1.*
11: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq2.*
12: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq3.*
13: /SID55/ptodata/1/pubpna/US10_NEW_PUB.seq4.*
14: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq.*
15: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
16: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
17: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq4.*
18: /SID55/ptodata/1/pubpna/US11_NEW_PUB.seq5.*
19: /SID55/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	785	98.2	1492	18	US-11-072-512-1704	Sequence 1704, Ap
2	775	97.0	497	11	US-10-475-075-564	Sequence 564, App
3	752	94.1	500	11	US-10-475-075-563	Sequence 563, App
4	223.5	28.0	1141	10	US-10-821-234-181	Sequence 181, App
5	115	14.4	159497	17	US-11-112-908-61	Sequence 61, Appl
6	115	14.4	171427	17	US-11-112-908-60	Sequence 60, Appl
7	85.5	10.7	16792	10	US-10-995-561-262	Sequence 262, App
8	85.5	10.7	17292	10	US-10-995-561-260	Sequence 260, App
9	85.5	10.7	17436	10	US-10-995-561-257	Sequence 257, App
10	85.5	10.7	17507	10	US-10-995-561-254	Sequence 254, App
11	85.5	10.7	17610	10	US-10-995-561-258	Sequence 258, App
12	85.5	10.7	17642	10	US-10-995-561-256	Sequence 256, App
13	85	10.6	3996	10	US-10-750-185-24510	Sequence 24510, A
14	85	10.6	3996	10	US-10-750-623-24510	Sequence 24510, A
15	84.5	10.6	8467	17	US-11-011-332A-83	Sequence 83, Appl
16	83.5	10.5	911	12	US-10-301-480-605341	Sequence 605341, A
17	83.5	10.5	911	12	US-10-301-480-610900	Sequence 610900, A
18	83.5	10.5	911	12	US-10-301-480-1218750	Sequence 1218750, A
19	83.5	10.5	911	12	US-10-301-480-1224309	Sequence 1224309, A
20	83	10.4	574	7	US-09-925-065A-467415	Sequence 467415, A
21	83	10.4	587	12	US-10-301-480-522972	Sequence 522972, A
22	83	10.4	587	12	US-10-301-480-1136381	Sequence 1136381, A
23	81.5	10.2	994	12	US-10-301-480-561980	Sequence 561980, A
24	81.5	10.2	994	12	US-10-301-480-1175389	Sequence 1175389, A
25	79.5	9.9	620	7	US-09-925-065A-691324	Sequence 691324, A
26	79.5	9.9	1311	17	US-11-128-061-1022	Sequence 1022, Ap
27	79.5	9.9	1311	17	US-11-128-061-4664	Sequence 4664, Ap
28	79.5	9.9	1311	17	US-11-128-049-1022	Sequence 1022, Ap
29	79.5	9.9	1311	17	US-11-128-049-4664	Sequence 4664, Ap
30	79	9.9	962	12	US-10-301-480-574920	Sequence 574920, A
31	79	9.9	962	12	US-10-301-480-1188329	Sequence 1188329, A
32	79	9.9	998	12	US-10-301-480-574919	Sequence 574919, A
33	79	9.9	998	12	US-10-301-480-1188328	Sequence 1188328, A
34	79	9.9	79329	11	US-10-330-773-483	Sequence 483, App
35	78.5	9.8	21293	11	US-10-330-773-272	Sequence 272, App
36	78.5	9.7	591	7	US-09-925-065A-898046	Sequence 898046, A
37	77.5	9.7	591	7	US-09-925-065A-13514	Sequence 13514, A
38	77.5	9.7	666	11	US-10-301-480-114751	Sequence 114751, A
39	77.5	9.7	666	12	US-10-301-480-728160	Sequence 728160, A
40	77.5	9.7	666	12	US-10-301-480-606658	Sequence 606658, A
41	77.5	9.7	989	12	US-10-301-480-1220067	Sequence 1220067, A
42	77.5	9.7	989	12	US-10-301-480-506732	Sequence 506732, A
43	77	9.6	603	7	US-09-925-065A-59108	Sequence 59108, A
44	77	9.6	1747	10	US-10-750-185-59108	Sequence 59108, A
45	77	9.6	1747	10	US-10-750-623-59108	Sequence 59108, A

ALIGNMENTS

RESULT 1
US-11-072-512-1704
; Sequence 1704, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKI
; APPLICANT: NAGAHARI, KENJI

```
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1704
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-1704

Alignment Scores:
Pred. No.: 2,19e-72 Length: 1492
Score: 785.00 Matches: 150
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 98.2% Indels: 0
DB: 18 Gaps: 0

US-10-664-025-3903 (1-153) x US-11-072-512-1704 (1-1492)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
Db 52 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTCTGAGAACCTC 111
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 112 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGCCCTCAGTTATCCAGTTGCAGAA 171
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 172 AGCAGAAATCTGTATATATTTGGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 231
QY 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValAlaAspCysGlyIleArgThr 80
Db 232 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 291
QY 81 ArgValValSerGluGluThrLeuPheGluLeuThrGluLeuTyrllePheThrProArgAsn 100
Db 292 AGGTTAGTTCTTGAGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 351
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
Db 352 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTCCACCTTAGGAAATCAGTGTGG 411
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db 412 CTTACACCAAGTTTCTACTCAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 471
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
Db 472 TTTCAGACAAACAGCAGAGAGTTAGGATTATTA 504
```

RESULT 2

```
US-10-475-075-564
; Sequence 564, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
```

```
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 564
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 52..495
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 52..102
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.699998092651
; OTHER INFORMATION: seq VMLLLAVLIWTGA/EN
US-10-475-075-564
```

```
Alignment Scores:
Pred. No.: 6,82e-72 Length: 497
Score: 775.00 Matches: 148
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.0% Indels: 0
DB: 11 Gaps: 0
```

US-10-664-025-3903 (1-153) x US-10-475-075-564 (1-497)

```
QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleTrpThrGlyAlaGluAsnLeu 20
Db 52 ATGGCGTTAGAAAGTCTTGATGCTCTCGCTGCTTGAATTTGGACCGGTCTGAGAACCTC 111
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 112 CATGTGAAATAAGTTGCTCTCTGGACTGGTTGATGGGTCTCAGTTATCCAGTTGCAGAA 171
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 172 AGCAGAAATCTGTATATATTTGGGATGAATTTACATCTGGGAATGGGCTGCCCTGCAAT 231
QY 61 ArgIleHisThrTyrlleValTyrlleGluPheIleTyrlleValAlaAspCysGlyIleArgThr 80
Db 232 CGGATACATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 291
QY 81 ArgValValSerGluGluThrLeuPheGluLeuThrGluLeuTyrllePheThrProArgAsn 100
Db 292 AGGTTAGTTCTTGAGAAATCTCTCTTTTCAACCGAGCTGTACTTTACCCCAAGGAAT 351
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleValTrp 120
Db 352 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTCCACCTTAGGAAATCAGTGTGG 411
QY 121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db 412 CTTACACCAAGTTTCTACTCAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 471
QY 141 PheGlnThrThrAlaGluLeuLeu 148
Db 472 TTTCAGACAAACAGCAGAGAGTTA 495
```

RESULT 3

```
US-10-475-075-563
; Sequence 563, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
```



```

; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 563
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 72..500
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 72..122
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10.6999998092651
; OTHER INFORMATION: seq VMLLAVLIWTGA/EN
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1..3
; OTHER INFORMATION: n = a, g, c or t
; US-10-475-075-563

Alignment Scores:
Pred. No.: 1,79e-69 Length: 500
Score: 752.00 Matches: 143
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 94.1% Indels: 0
DB: 11 Gaps: 0

US-10-664-025-3903 (1-153) x US-10-475-075-563 (1-500)

QY 1 MetAlaLeuGluValLeuMetLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeu 20
Db 72 ATGGCGTTAGAGTCTTGATGCTCTCGCTGCTTGATTGGACCGGTGCTGAGAACCTC 131
QY 21 HisValLysLeuSerCysSerLeuAspTrpLeuMetValSerValLeuProValAlaGlu 40
Db 132 CATGTGAAAATAAGTTGCTCTCGGACTGGTTCATGGTCTCAGTTATCCAGGTGCAGAA 191
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 192 AGCAGAAATCTGTATATATTTGCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAA 251
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 252 CGGATACATACATATGATATATGATTTATATATCTTTGTTGATGTGGCATCAGGACA 311
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db 312 AGGGTAGTTCTTCAGGAAATCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGGAT 371
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Db 372 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 431
QY 121 LeuThrProValSerThrGluAsnGluIleLeuLeuAspProSerProPheIleAlaAsp 140
Db 432 CTTACACCAAGTTCTTACTGAGAAATGAATAAAATGGATCGCTAGTCTCTTTATTTGCTG 491
QY 141 PheGlnThr 143
Db 492 TTTTCAGACA 500

RESULT 4
US-10-821-234-181
; Sequence 181, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 181
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-821-234-181

Alignment Scores:
Pred. No.: 1,46e-13 Length: 1141
Score: 223.50 Matches: 40
Percent Similarity: 60.0% Conservative: 32
Best Local Similarity: 33.3% Mismatches: 43
Query Match: 28.0% Indels: 5
DB: 10 Gaps: 2

US-10-664-025-3903 (1-153) x US-10-821-234-181 (1-1141)

QY 5 ValLeuMetLeuAlaValLeuLeuTrpThrGlyAlaGluAsnLeuHisValLysLeu 24
Db 316 CTGATGATCTCTCTCACCTCTGGGTTTTCAGCGGTTTCAGGACAAAGTCCATGACTGTG 375
QY 25 SerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeu 44
Db 376 CTGTGCTCCATAGACTGTTTCATGGTGCACAGTGCACCTTTCCTAAACAACGATGTG 435
QY 45 TyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThr 64
Db 436 TGTGTACACTTTCATGAACATACACTTGGGCTGGGTTGCCCCCAACCAACCATGTTTCAGCCA 495
QY 65 TyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSer 84
Db 496 CAGCGCTACAGTTACCTACCTGTTACTGATGTGGCATCAGGGCCCAAGGTGTCTCT 555
QY 85 GluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAsp 104
Db 556 CAGGACATGTTTATCTACAGCACTGAGATACACTACTCTTCTTAAGGCGC-----ACG 606
QY 105 ProGlnGlu-----IleHisLeuGluCysSerThrSerArgLysSerValTrpLeuThr 122
Db 607 CCATCTAAGTTTGTGATCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666

RESULT 5
US-11-112-908-61/c
; Sequence 61, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Davis, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61

```

```
; LENGTH: 159497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-61

Alignment Scores:
Pred. No.: 11.5 Length: 159497
Score: 115.00 Matches: 45
Percent Similarity: 46.5% Conservative: 22
Best Local Similarity: 31.2% Mismatches: 59
Query Match: 14.4% Indels: 19
DB: 17 Gaps: 6

US-10-664-025-3903 (1-153) x US-11-112-908-61 (1-159497)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuLeuThrGlyAlaGluAsnLeu 20
DB 147611 TTGCGGTTAAATGGGCTCTTTTCTCAATTCCTTGATGACCTGTGCTGGGACTGG 147552
QY 21 His---ValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAla 39
DB 147551 TCGGCTATTCAAGTACAATGCACCCAGTTTGTGTCGCCAGGATTAAACCCACGATA 147492
QY 40 GluSerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla 59
DB 147491 ---TTTCACAAATTGATATGAACCCCGATGAAGTGTCTTAGGAGATGGCTGCCT-TGT 147436
QY 60 AsnArgIleHisThrTyrVal---TyrGluPheIleTyrLeuValArgAspCysGlyIle 78
DB 147435 AACCTAGTTTTCGCCAAATATCTACTATGAGTTTCTACCATCTCCATGACTGTGGTAAC 147376
QY 79 ArgThrArgValValSer-GluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrPr 98
DB 147375 TCACCC-----TCTCCAGGAAGTTCTTCTGCTTAAACTAAATCAAGTATATCTC 147325
QY 98 oArgAsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSe 118
DB 147324 AAGA-----GACTACTCCGATCTGAAATGCCTCTGTCTGTTCGCCCAATCAGAA 147271
QY 118 rValTrpLeuThrProValSer-----ThrGluAsnG1 129
DB 147270 TCCTCTTTTAAATGTAGTTGAAATGAGAGATAAGGAAACTGACAATGTTACTGAATGGGA 147211
QY 129 uIleLysLeu 132
DB 147210 GATAGAGGTG 147201

RESULT 6
US-11-112-908-60/c
; Sequence 60, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60
; LENGTH: 171427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-60

Alignment Scores:
Pred. No.: 1.04e+03 Length: 16792
Score: 85.50 Matches: 36
Percent Similarity: 41.5% Conservative: 23
Best Local Similarity: 25.4% Mismatches: 58
Query Match: 10.7% Indels: 25
DB: 10 Gaps: 5

US-10-664-025-3903 (1-153) x US-10-995-561-262 (1-16792)

Alignment Scores:
Pred. No.: 12.6 Length: 171427
Score: 115.00 Matches: 45
Percent Similarity: 46.5% Conservative: 22
Best Local Similarity: 31.2% Mismatches: 59
Query Match: 14.4% Indels: 19
DB: 17 Gaps: 6

US-10-664-025-3903 (1-153) x US-11-112-908-60 (1-171427)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuLeuThrGlyAlaGluAsnLeu 20
DB 53747 TTGCGGTTAAATGGGCTCTTTTCTCAATTCCTTGATGACCTGTGCTGGGACTGG 53688
QY 21 His---ValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAla 39
DB 53687 TCGGCTATTCAAGTACAATGCACCCAGTTTGTGTCGCCAGGATTAAACCCACGATA 53628
QY 40 GluSerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla 59
DB 53627 ---TTTCACAAATTGATATGAACCCCGATGAAGTGTCTTAGGAGATGGCTGCCT-TGT 53572
QY 60 AsnArgIleHisThrTyrVal---TyrGluPheIleTyrLeuValArgAspCysGlyIle 78
DB 53571 AACCTAGTTTTCGCCAAATATCTACTATGAGTTTCTACCATCTCCATGACTGTGGTAAC 53512
QY 79 ArgThrArgValValSer-GluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrPr 98
DB 53511 TCACCC-----TCTCCAGGAAGTTCTTCTGCTTAAACTAAATCAAGTATATCTC 53461
QY 98 oArgAsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSe 118
DB 53460 AAGA-----GACTACTCCGATCTGAAATGCCTCTGTCTGTTCGCCCAATCAGAA 53407
QY 118 rValTrpLeuThrProValSer-----ThrGluAsnG1 129
DB 53406 TCCTCTTTTAAATGTAGTTGAAATGAGAGATAAGGAAACTGACAATGTTACTGAATGGGA 53347
QY 129 uIleLysLeu 132
DB 53346 GATAGAGGTG 53337

RESULT 7
US-10-995-561-262
; Sequence 262, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 262
; LENGTH: 16792
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-262

Alignment Scores:
Pred. No.: 1.04e+03 Length: 16792
Score: 85.50 Matches: 36
Percent Similarity: 41.5% Conservative: 23
Best Local Similarity: 25.4% Mismatches: 58
Query Match: 10.7% Indels: 25
DB: 10 Gaps: 5

US-10-664-025-3903 (1-153) x US-10-995-561-262 (1-16792)
```

```
QY 3 LeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAenLeuHisVal 22
Db 1276 CTGGAATTGCTGTACAGATGCAACAAATCCAGAAATGGTCTTGAACCTGTGAAGAA 1335
QY 23 LysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArg 42
Db 1336 AAACGTGACACTAGCTAAGAAATACACTGCAG-----1365
QY 43 AsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla-----59
Db 1366 -----GCTGATGCTGCTACCTTGAATTCAGGACAAACCGGTACAAATGTGAG 1410
QY 60 AsnArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArg 79
Db 1411 TCAGATGTCATTATGATCAGAGTGTGAAGTCTCATCAGGAGCTGCAGGTGGAT 1470
QY 80 ThrArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArg 99
Db 1471 CTCAGATCTCGGGATGAGAAATTACTACCAGCTAGAGAGCTGGCTTTTAGGGTCATG 1530
QY 100 AsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThr---SerArgLys---117
Db 1531 CGTCTTCAGATGAGCTGGTCACCTTGCCTAGAGTGTACAAACCTGTACCGGAAGGGT 1590
QY 118 -----SerValTrpLeuThrProValSer-----ThrGluAenGluIleLysLeu 132
Db 1591 CATTTCACTTCATTGAATTTGGTTCCACCCTCTACTTTTACCACCACTCATCTGAAAGCA 1650
QY 133 AspPro 134
Db 1651 GAACCC 1656

RESULT 8
US-10-995-561-260
; Sequence 260, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260
; LENGTH: 17292
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-260
```

```
Alignment Scores:
Pred. No.: 1.08e+03 Length: 17292
Score: 85.50 Matches: 36
Percent Similarity: 41.5% Conservative: 23
Best Local Similarity: 25.4% Mismatches: 58
Query Match: 10.7% Indels: 25
DB: 10 Gaps: 5

US-10-664-025-3903 (1-153) x US-10-995-561-260 (1-17292)
```

```
QY 3 LeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAenLeuHisVal 22
Db 1069 CTGGAATTGCTGTACAGATGCAACAAATCCAGAAATGGTCTTGAACCTGTGAAGAA 1128
QY 23 LysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArg 42
Db 1129 AAACGTGACACTAGCTAAGAAATACACTGCAG-----1158
QY 43 AsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla-----59
Db 1159 -----GCTGATGCTGCTACCTTGAATTCAGGACAAACCGGTACAAATGTGAG 1203
```

```
QY 60 AsnArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArg 79
Db 1204 TCAGATGTCATTATGATCAGAGTGTGAAGTCTCATCAGGAGCTGCAGGTGGAT 1263
QY 80 ThrArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArg 99
Db 1264 CTCAGATCTCGGGATGAGAAATTACTACCAGCTAGAGAGCTGGCTTTTAGGGTCATG 1323
QY 100 AsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThr---SerArgLys---117
Db 1324 CGTCTTCAGATGAGCTGGTCACCTTGCCTAGAGTGTACAAACCTGTACCGGAAGGGT 1383
QY 118 -----SerValTrpLeuThrProValSer-----ThrGluAenGluIleLysLeu 132
Db 1384 CATTTCACTTCATTGAATTTGGTTCCACCCTCTACTTTTACCACCACTCATCTGAAAGCA 1443
QY 133 AspPro 134
Db 1444 GAACCC 1449

RESULT 9
US-10-995-561-257
; Sequence 257, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257
; LENGTH: 17436
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-257
```

```
Alignment Scores:
Pred. No.: 1.09e+03 Length: 17436
Score: 85.50 Matches: 36
Percent Similarity: 41.5% Conservative: 23
Best Local Similarity: 25.4% Mismatches: 58
Query Match: 10.7% Indels: 25
DB: 10 Gaps: 5

US-10-664-025-3903 (1-153) x US-10-995-561-257 (1-17436)
```

```
QY 3 LeuGluValLeuMetLeuLeuAlaValLeuIleTrpThrGlyAlaGluAenLeuHisVal 22
Db 1276 CTGGAATTGCTGTACAGATGCAACAAATCCAGAAATGGTCTTGAACCTGTGAAGAA 1335
QY 23 LysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArg 42
Db 1336 AAACGTGACACTAGCTAAGAAATACACTGCAG-----1365
QY 43 AsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla-----59
Db 1366 -----GCTGATGCTGCTACCTTGAATTCAGGACAAACCGGTACAAATGTGAG 1410
QY 60 AsnArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArg 79
Db 1411 TCAGATGTCATTATGATCAGAGTGTGAAGTCTCATCAGGAGCTGCAGGTGGAT 1470
QY 80 ThrArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArg 99
Db 1471 CTCAGATCTCGGGATGAGAAATTACTACCAGCTAGAGAGCTGGCTTTTAGGGTCATG 1530
QY 100 AsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThr---SerArgLys---117
Db 1159 -----GCTGATGCTGCTACCTTGAATTCAGGACAAACCGGTACAAATGTGAG 1203
```



```
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 256
; LENGTH: 17642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-256

Alignment Scores:
Pred. No.: 1.1e+03      Length: 17642
Score: 85.50           Matches: 36
Percent Similarity: 41.5%      Conservative: 23
Best Local Similarity: 25.4%    Mismatches: 58
Query Match: 10.7%           Indels: 25
DB: 10                  Gaps: 5

US-10-664-025-3903 (1-153) x US-10-995-561-256 (1-17642)

QY 3 LeuGluValLeuMetLeuLeuAlaValLeuIleThrGlyAlaGluAsnLeuHieVal 22
Db 1276 CTGGATTGCTGTACAGATTGCAACAAATCCAGATGGTCTTTGAAGAA 1335
QY 23 LyeIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArg 42
Db 1336 AAACAGACACTAGCTAAGAAATACACTGCAG----- 1365
QY 43 AsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla----- 59
Db 1366 -----GCTGATGCTGTCTACCTGGAATCAGGACAAACCGGTACAATGTGAG 1410
QY 60 AsnArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArg 79
Db 1411 TCAGATGTCATTATGATTCAGAGGTGTGAAGGCTCTATCAGGCGAGCTGCAGGTGGAT 1470
QY 80 ThrArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArg 99
Db 1471 CTCAGATCTCGGGATGAGATTACTACCAGCTAGAAGAGCTGGCTTTAGGTCATG 1530
QY 100 AsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThr---SerArgLys--- 117
Db 1531 CGTCTTCAGATGAGCTGGTGCACCTTGGCGTCTAGAGGTGTACAAACCTGTACCGGAAGGT 1590
QY 118 -----SerValTrpLeuThrProValSer-----ThrGluAsnGluIleLysLeu 132
Db 1591 CATTCACCTCACTTGAATGGTTCCACCCTCTACTTTAACCAACCACCTCATCTGAAAGCA 1650
QY 133 AspPro 134
Db 1651 GAACCC 1656

RESULT 13
US-10-750-185-24510
; Sequence 24510, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 24510
; LENGTH: 3996
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-623-24510

Alignment Scores:
Pred. No.: 220          Length: 3996
Score: 85.00           Matches: 33
Percent Similarity: 41.4%      Conservative: 22
Best Local Similarity: 24.8%    Mismatches: 60
Query Match: 10.6%           Indels: 18
DB: 4                  Gaps: 4

US-10-664-025-3903 (1-153) x US-10-750-623-24510 (1-3996)

QY 9 LeuAlaValLeuIleTrpThrGlyAlaGluAsnLeuHisValLysIleSerCysSerLeu 28
Db 466 CTGATTTTATTATTTATGAAGCCATATCTGCTTGACATTCAGGGGTATCATGTTTCTT 525
QY 29 AspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeuTyrIlePheAla 48
Db 526 CACTGGTCCACACACACCCCGGACCCAAAGGAAA-----GCT 564
QY 49 Asp-----GluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThrTyr 65
Db 565 GATGCCACGACGCGCTGCACATACTTACCACATCCCTGGCAAGGGCTGCAT----- 618
QY 66 ValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSerGlu 85
Db 619 -----CTGTTTCAGAGAGAGGGGTCTCTTAATTGCGACAAAGATGCTGCTGCTT 666
QY 86 GluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAspPro 105
Db 667 GAACATATGAAGTCTCTGAGACATTATATTGAGACCT-----ATTGACTTGTCTCTCC 720
QY 106 GlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrpLeuThrProValSer 125
Db 721 CAGGTAGATATAATAAGTCCATGAGAGAGGAGGACTGTGTGGGAAGACTCCGCTGTC 780
QY 126 ThrGluAsnGluIleLysLeuAspProSerProPheIle 138
Db 781 TCTCCAGGTTCATCTGGCCACCCACCCTCTCTCCCTACCTA 819

RESULT 14
US-10-750-623-24510
; Sequence 24510, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 24510
; LENGTH: 3996
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-623-24510

Alignment Scores:
Pred. No.: 220          Length: 3996
Score: 85.00           Matches: 33
Percent Similarity: 41.4%      Conservative: 22
Best Local Similarity: 24.8%    Mismatches: 60
Query Match: 10.6%           Indels: 18
DB: 4                  Gaps: 4
```


GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2006, 13:30:21 ; Search time 5379.3 Seconds
(without alignments)
1996.103 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLIWTGAENL.....PSFPIADFTQTAELGLLIF 153

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/abs/ABSSWEB spool/US10664025/runat_08052006_173646_29404/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=dct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pf0 -NORM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes07
-USER=US10664025 @CGN_1_1_6731@runat_08052006_173646_29404 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsl:*
10: gb_gsl2:*
11: gb_gsl3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	789	98.7	626	BI463012	BI463012 603204535
2	789	98.7	776	BG772527	BG772527 602720619
3	789	98.7	878	BI520154	BI520154 603071295
4	767	96.0	577	BI118111	BI118111 603118111
5	743	93.0	778	BI459538	BI459538 603200545
6	706	88.4	619	AW664990	AW664990 bi98h06.x
7	670	83.9	516	BE972717	BE972717 601652136

C	8	563.5	70.5	508	1	AA572898	hfl7c01.x
C	9	538	67.3	430	1	AA608835	af03h06.e
C	10	514.5	64.4	480	1	AI554902	tes3903.x
	11	431.5	54.0	598	2	BG081002	BG081002 H3059H12-
	12	431.5	54.0	629	5	BY731676	BY731676
	13	431.5	54.0	862	7	CO810251	AGENCOURT
C	14	429	53.7	331	1	AI113881	GD59C04.x
	15	395.5	49.5	642	2	BB557840	BB557840 BB557840
C	16	329	41.2	321	1	AI028204	ov96b07.x
C	17	288	36.0	496	2	BG067944	H3059H12-
C	18	285.5	35.7	587	5	C88054	C88054 C88054 Mous
C	19	280	35.0	504	7	CO798599	AGENCOURT
C	20	276.5	34.6	576	1	AU045772	AU045772
C	21	235.5	29.5	431	2	BF708434	MI-P-AYO-
C	22	234.5	29.3	497	2	BF709198	BF709198 MI-P-AYO-
	23	234	29.3	598	2	BG077467	H3015G03-
	24	234	29.3	663	7	CK334352	H3015G03-
	25	234	29.3	700	5	BY705703	BY705703
	26	234	29.3	785	7	CK128283	AGENCOURT
	27	234	29.3	1071	4	AK005501	Mus muscu
	28	233.5	29.2	668	5	BY736907	BY736907
	29	230.5	28.8	687	6	CF367041	841759 MA
	30	229.5	28.7	655	6	CF362621	830339 MA
	31	227	28.4	666	2	BG620807	BG620807 602617729
C	32	225.5	28.2	378	7	CO940089	UWC-pd45f
C	33	224	28.0	868	7	CK769542	957583 MA
	34	223.5	28.0	696	6	CB996101	AGENCOURT
	35	223.5	28.0	820	6	CF593309	CF593309
	36	223.5	28.0	882	2	BG199300	RST18581
	37	223.5	28.0	883	5	BX386304	BX386304
	38	223.5	28.0	895	5	CA489318	AGENCOURT
	39	223.5	28.0	943	4	CR604394	full-leng
	40	223.5	28.0	965	5	BX359837	BX359837
	41	223.5	28.0	1056	4	CR618096	full-leng
	42	223.5	28.0	1075	4	CR603619	full-leng
	43	220.5	27.6	812	2	BG202365	RST21722
	44	219.5	27.5	168	6	CB298889	220023 re
	45	219.5	27.5	752	5	BX110037	BX110037

ALIGNMENTS

RESULT 1

BI463012

LOCUS

DEFINITION

603204535F1 NIH_MGC_97 Homo sapiens

mRNA sequence.

ACCESSION

BI463012

VERSION

BI463012.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 626)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Cloned through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11682 row: e column: 09

High quality sequence stop: 624.

Location/Qualifiers

1..626

source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5270384"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,33e-86 Length: 626
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 3 Gaps: 0

US-10-664-025-3903 (1-153) x BI463012 (1-626)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTTrpThrGlyAlaGluAenLeu 20
Db 57 ATGGCGTTAGAAAGTCTTGATCTCTCGTCTTGTGATTTGGACCGGTCTGAGAACCTC 116
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 117 CATGTGAAATAAGTTCCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 176
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 177 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 236
QY 61 ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr 80
Db 237 CGGATACATACATATGATGATGATATATATCTTGTTCGTGATTTGGCATCAGGACA 296
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
Db 297 AGGGTAGTTTCTGAGAAACTCTCCTTTTCAACCGAGCTGACTTTACCCCAAGGAAT 356
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Db 357 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 416
QY 121 LeuThrProValSerThrGluAenGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db 417 CTTACACCAAGTTCTTACTGAGAATGAAATATAAATTTGGATCTGATCTCTTTATGCTGAC 476
QY 141 PheGlnThrThrAlaGluLeuGlyLeuLeu 151
Db 477 TTTGAGAACACGACAGAGTTAGGATTATTA 509
```

RESULT 2

```
BG772527
LOCUS BG772527 776 bp mRNA linear EST 15-MAY-2001
DEFINITION 602720619F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837612 5',
mRNA sequence.
```

```
ACCESSION BG772527
```

```
VERSION BG772527.1 GI:14083180
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
ORGANISM
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
```

REFERENCE

```
1 (bases 1 to 776)
NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsbbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10770 row: e column: 05
High quality sequence stop: 773.
```

FEATURES

source

```
1..776
Location/Qualifiers
1..776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4837612"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 3.11e-86 Length: 776
Score: 789.00 Matches: 151
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 2 Gaps: 0

US-10-664-025-3903 (1-153) x BG772527 (1-776)

QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTTrpThrGlyAlaGluAenLeu 20
Db 57 ATGGCGTTAGAAAGTCTTGATCTCTCGTCTTGTGATTTGGACCGGTCTGAGAACCTC 116
QY 21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 117 CATGTGAAATAAGTTCCTCTGGACTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 176
QY 41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 177 AGCAGAAATCTGTATATATTTGGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAT 236
QY 61 ArgIleHisThrTyrrValTyrrGluPheIleTyrrLeuValArgAspCysGlyIleArgThr 80
Db 237 CGGATACATACATATGATGATGATATATCTTGTTCGTGATTTGGCATCAGGACA 296
QY 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrrPheThrProArgAsn 100
Db 297 AGGGTAGTTTCTGAGAAACTCTCCTTTTCAACCGAGCTGACTTTACCCCAAGGAAT 356
QY 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrp 120
Db 357 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 416
QY 121 LeuThrProValSerThrGluAenGluIleLysLeuAspProSerProPheIleAlaAsp 140
Db 417 CTTACACCAAGTTCTTACTGAGAATGAAATATAAATTTGGATCTGATCTCTTTATGCTGAC 476
```



```

QY      141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
      |||||||
Db      477 TTTCAGACAACAGCAGAAGAGTTAGGATTATTA 509

RESULT 3
BI520154
LOCUS   603071295F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163455 5',
DEFINITION mRNA sequence.
ACCESSION BI520154
VERSION   1 GI:15344946
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 878)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabs-r@mail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: L1A11405 Row: m Column: 24
          High quality sequence stop: 851.
          Location/Qualifiers
            1..878
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5163455"
              /tissue_type="medulla"
              /lab_host="DH10B"
              /clone_lib="NIH MGC 119"
              /notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: EcoRV (destroyed); RNA source normal medulla from
              anonymous male age 27. Library is oligo-dT primed and
              directionally cloned (ScorV site is destroyed upon
              cloning). Average insert size 1.3 kb, insert size range
              0.9-3 kb. Library is normalized and enriched for
              full-length clones and was constructed by C. Gruber
              (Invitrogen). Research Genetics tracking code 013. Note:
              this is a NIH_MGC Library."

ORIGIN
source
Alignment Scores:
Pred. No.:      3 67e-86      Length:      878
Score:          789.00      Matches:      151
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    98.7%      Indels:      0
DB:             Gaps:      0

US-10-664-025-3903 (1-153) x BI520154 (1-878)

QY      1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuThrThrGlyAlaGluAsnLeu 20
      |||||||
Db      30 ATGGCGTTAGAGTCTTGATGCTCTCGTGTCTTGATTTGACCGGTCTGAGACCTC 89

QY      21 HisValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
      |||||||
Db      90 CATGTGAAATAAGTTGCTCTGGAAGTCTGATGCTCTGATGCTCTGATGCTCTGAGAA 149

QY      41 SerArgAsnLeuTyrllePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
      |||||||

```

```

Db      150 AGCAGAAATCTGTATATATTTTGGGATGATTTACATCTGGAAATGGCTGCCCTGCAAT 209
QY      61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
      |||||||
Db      210 CGGATACATACATATGTATATGAGTTTATATATCTTGTTCGTGATTTGGCATCAGGACA 269

QY      81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
      |||||||
Db      270 AGGGTAGTTTCTGAGGAAACTCTCCCTTTTCAACACGAGCTGTACTTTACCCCAAGGAAT 329

QY      101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr 120
      |||||||
Db      330 ATAGATCATGACCCCTCAGGAATCCATTGGAGTGTTCACCTCTAGGAATCAGTGTGG 389

QY      121 LeuThrProValSerThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAsp 140
      |||||||
Db      390 CTTACACCAAGTTTCTACTGAGAATGAAATAAAATTTGGATCCTAGTCTCTTTTATTGCTGAC 449

QY      141 PheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
      |||||||
Db      450 TTTCAGACAACAGCAGAAGAGTTAGGATTATTA 482

RESULT 4
BI118111
LOCUS   EX118111 Soares testis NHT Homo sapiens cDNA clone IMAGE:9998D122574
DEFINITION ; IMAGE:1030619, mRNA sequence.
ACCESSION EX118111
VERSION   EX118111.1 GI:27881212
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
          Radelof,U., Schneider,D. and Korn,B.
TITLE     Human Unigeneset - RZPD3
JOURNAL   Unpublished (2003)
COMMENT   Contact: Ina Rolfes
          RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
          Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
          RZPD; IMAGE:9998D122574.
          RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
          Human Unigeneset - RZPD3 (RZPDLIB No.972)
          http://www.rzpd.de/CloneCards/cgi-
          bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfes
          RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
          Heubnerweg 6, D-14059 Berlin, Germany
          Tel: +49 30 32639 101
          Fax: +49 30 32639 111
          www.rzpd.de
          This clone is available royalty-free from RZPD;
          contact RZPD (clone@rzpd.de) for further information. Seq primer:
          M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.

FEATURES
Location/Qualifiers
  1..577
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:9998D122574 ; IMAGE:1030619"
    /sex="male"
    /lab_host="DH10B"
    /notes="Vector: p7T73D-Pac (Pharmacia) with a modified
    polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
    was prepared from mRNA obtained from Clontech
    Laboratories, Inc., and primed with a Not I - oligo(dT)
    primer [5'
    TGTTCACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTTTT 3'].
    Double-stranded cDNA was ligated to Eco RI adaptors
    (Pharmacia), digested with Not I and cloned into the Not I

```

and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 1.07e-83 Length: 577
Score: 767.00 Matches: 146
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.0% Indels: 0
DB: 5 Gaps: 0

US-10-664-025-3903 (1-153) x BX118111 (1-577)

Qy 6 LeuMetLeuAlaValLeuLeuThrGlyAlaGluAsnLeuHisVallylsleSer 25
Db 3 TTGATGCTCCTCGCTGCTGTGATTTGGACCGTGTGAGAACCTCCATGTGAAATAAGT 62
Qy 26 CysSerLeuAspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeuTyr 45
Db 63 TGCTCTCGACTGGTGTGATGCTCAGTTATCCCGTTGCAGAAAGCAGAAATCTGTAT 122
Qy 46 IlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThrTyr 65
Db 123 ATATTGCGGATGAATTACATCTGGGATGGCTGCCCTGCAATCGGATACATACATAT 182
Qy 66 ValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSerGlu 85
Db 183 GTATATGAGTTTATATATCTTGTGATGTTGGCATCAGGACAGGGTAGTTCTGAG 242
Qy 86 GluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAspPro 105
Db 243 GAAACTCTCTTTTCAAACCGAGCTGTACTTTACCCGAGGAATATAGATCATGACCT 302
Qy 106 GlnGluIleHisLeuGluCysSerThrSerArgLysSerValTrpLeuThrProValSer 125
Db 303 CAGGAAATCCATTGGAGTGTCCACCTCTAGGAAATCAGTGTGGCTTACACAGTTTCT 362
Qy 126 ThrGluAsnGluIleLysLeuAspProSerProPheIleAlaAspPheGlnThrAla 145
Db 363 ACTGAGAATGAATAAAATGGATCCTAGTCTCTTTTATTGCTGACTTTTCAGACACAGCA 422
Qy 146 GluGluLeuGlyLeuLeu 151
Db 423 GAAGAGTTAGGATTAATA 440

RESULT 5

BI459538

LOCUS 603200545F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5266642 5',
DEFINITION mRNA sequence.

ACCESSION

BI459538

VERSION

BI459538.1 GI:15250194

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>
Plate: LLAM11672 row: i column: 11
High quality sequence stop: 674.

FEATURES

source

1..778
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5266642"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1.44e-80 Length: 778
Score: 743.00 Matches: 148
Percent Similarity: 96.8% Conservative: 2
Best Local Similarity: 95.5% Mismatches: 2
Query Match: 93.0% Indels: 3
DB: 3 Gaps: 0

US-10-664-025-3903 (1-153) x BI459538 (1-778)

Qy 1 MetAlaLeuGluValLeuMetLeuAlaValLeuIleThrGlyAlaGluAsnLeu 20
Db 191 ATGGCGTTAGAGTCTTATGCTCTCTGCTGCTTTGATTGGACCGTGCTGAGAACCTC 250
Qy 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 251 CATGTGAANAATAGTTGCTCTCTGACGTGGTGTGATGGTCTCAGTTATCCAGTTGCAGAA 310
Qy 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 311 AGCAGAAATCTGATATATTTGCGGATGAATTACATCTGGGAATGGGCTGCCCTGCAAA 370
Qy 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValAlaAspCysGlyIleArgThr 80
Db 371 CGGATACATACATATGATATGAGTTATATATCTTTGTTGGATTGTGGCATCAGGACA 430
Qy 81 ArgValValSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArgAsn 100
Db 431 AGGGTAGTTTCTGAGGAAATCTCTCTTTTCAACCGAGCTACTTTACCCCAAGAA 490
Qy 101 IleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgLysSerValTyr 120
Db 491 ATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGGAAATCAGTGTGG 550
Qy 121 LeuThrProValSerThrGluAsnGlu--IleLysLeuAspProSerProPheIleAla 140
Db 551 CTTACACCAAGTTTCTACTGAGAACTGAAACTACAAATGGATCTAGTCTCTTTTATGCTG 610
Qy 140 spPheGlnThrThrAla-GluGluLeuGlyLeuLeuIle 152
Db 611 ACTTTCAGACACACAGCCAGAGATTACCGATTATATT 649

RESULT 6

AW664990/c

LOCUS

DEFINITION hi98h05.x1 Soares NFL T.GBC S1 Homo sapiens cDNA clone
IMAGE:2980379 3', mRNA sequence.

ACCESSION

AW664990

VERSION

AW664990.1 GI:7457534

KEYWORDS

EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 619)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 460.
Location/Qualifiers
1..619
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2980379"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
note="Organ: pooled; Vector: pVT3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbH19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.: 3..81e-76 Length: 619
Score: 706.00 Matches: 139
Percent Similarity: 97.2% Conservative: 1
Best Local Similarity: 96.5% Mismatches: 4
Query Match: 88.4% Indels: 1
DB: 1 Gaps: 0
US-10-664-025-3903 (1-153) x AW664990 (1-619)
QY 8 LeuLeuAlaValLeuIleThrPheGlyAlaGluAsnLeuHisVallylsleSerCysSer 27
Db 613 GTCTCGCTGCTTGATTGG-ACCGGTGCTGAGACCTCCATGTGAAATATAGTTGCTCT 555
QY 28 LeuAspTrpLeuMetValSerValIleProValAlaGluSerArgAsnLeuTyrIlePhe 47
Db 554 CTGACTTGGTTCATGGTCTCAGTTATCCAGTTGCGAANGCAGAAATCTGTATATATT 495
QY 48 AlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThrTyrValTyr 67
Db 494 GCGGATGAATTACATCTGGGATGGCTGCCCTGCCAATCGCATACATATGATATAT 435
QY 68 GluPheIleTyrLeuValArgAspCysGlyIleArgThrArgValValSerGluIleThr 87
Db 434 GAGTTATATATCTTGTTCGTGATTGTGGCATCAGACAAAGGGTAGTTTCTGAGGAACT 375
QY 88 LeuLeuPheGlnThrGluLeuTyrPheThrProArgAsnIleAspHisAspProGlnGlu 107
Db 374 CTCCTTTTCAACCGAGCTGTACTTTACCCAGAGATATAGATCATGACCTCAGGAA 315
QY 108 IleHisLeuGluCysSerThrSerArgIysSerValTrpLeuThrProValSerThrGlu 127
Db 314 ATCCATTGGAGTGTCCACCTCTAGGAATCAGTGTGCTTACACCACTTCTACTGAG 255
QY 128 AsnGluIleLysLeuAspProSerProPheIleAlaAspPheGlnThrThrAlaGlu 147

Db 254 AATGAATAAATTGGATCTAGTCTCTTTATTGCTGACTTTTCAGACAAACAGAGAAG 195
QY 148 LeuGlyLeuLeu 151
Db 194 TTAGGATTATTA 183
RESULT 7
BE972717 516 bp mRNA linear EST 04-OCT-2000
LOCUS 601652136F1 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:3935288 5',
DEFINITION mRNA sequence.
ACCESSION BE972717
VERSION BE972717.1 GI:10586053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCW777 row: h column: 09
High quality sequence stop: 505.
Location/Qualifiers
1..516
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3935288"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_82"
note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccattggcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCGCATTTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
ORIGIN
Alignment Scores:
Pred. No.: 8..07e-72 Length: 516
Score: 670.00 Matches: 130
Percent Similarity: 98.5% Conservative: 1
Best Local Similarity: 97.7% Mismatches: 1
Query Match: 83.9% Indels: 1
DB: 2 Gaps: 0
US-10-664-025-3903 (1-153) x BE972717 (1-516)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuIleThrThrGlyAlaGluAsnLeu 20
Db 120 ATGGCGTTAGAAGTCTTGATGCTCTCGCTGCTTTGATTGGACCGGTCTGAGAACCCTC 179
QY 21 HisVallylsleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 180 CATGTGAANAATAGTTGCTCTCTGACTGGTGTGATGGTCTCAGTTATCCAGTTGAGAA 239

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pVT33 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Alignment Scores:

Pred. No.: 1,13e-55 Length: 430
Score: 538.00 Matches: 101
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 67.3% Indels: 0
DB: 1 Gaps: 0

US-10-664-025-3903 (1-153) x AA608835 (1-430)

Qy 51 LeuHisLeuGlyMetGlyCysProAlaAsnArgIleHisThrTyrValTyrGluPheIle 70
430 TTACATCTGGGAATGGCTGCCCTCGCAATCGGATACATATGATATGAGTTTATA 371
Qy 71 TyrLeuValArgAspCysGlyIleArgThrArgValValSerGluGluThrLeuLeuPhe 90
370 TATCTTCTTCGATCTGGCATCAGACCAAGGTAGTTTCTGAGGAACTCTCTCTTTT 311
Qy 91 GluThrGluLeuTyrPheThrProArgAsnIleAspHisPProGlnGluIleHisLeu 110
310 CAACCCGAGCTGTACTTTTACCCCAAGGAATATAGATCATGACCTCAGGAAATCCATTTG 251
Qy 111 GluCysSerThrSerArgLysSerValTyrLeuThrProValSerThrGluAsnGluIle 130
250 GAGTGTTCACCTCTAGGAATCAGTGTGCTTACACAGTTTCTACTGAGATGAATA 191
Qy 131 LysLeuAspProSerProPheIleAlaAspPheGlnThrThrAlaGluGluLeuGlyLeu 150
190 AAATTGCATCTAGTCTCTTTTATTGCTGACTTTTCAGACACAGCAGAGAGTTAGGATTA 131
Qy 151 Leu 151
130 TTA 128

RESULT 10

AI554902/c
LOCUS
DEFINITION
IMAGE:2090452 3', mRNA sequence.

AI554902
VERSION
KEYWORDS
SOURCE

AI554902.1 GI:4487265
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 480)

NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 656 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 458.

Location/Qualifiers

1..480

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2090452"

/lab_host="DH10B"

FEATURES

source

RESULT 11

BG081002

LOCUS

DEFINITION

H3059H12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone

H3059H12 5', mRNA sequence.

ACCESSION

BG081002

VERSION

BG081002.2 GI:40015217

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 598)

AUTHORS

Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H. III, Becker,K.G. and Ko,M.S.H.

TITLE

Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

PUBMED

10922068

COMMENT

On Jan 26, 2001 this sequence version replaced gi:12563570.

Other_ESTs: H3059H12-3

Contact: George J. Kargul

/clone lib="Soares_NFL_T_GBC_S1"

/notes="Organ: pooled; Vector: pVT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 1,02e-52 Length: 480
Score: 514.50 Matches: 97
Percent Similarity: 92.6% Conservative: 3
Best Local Similarity: 89.8% Mismatches: 1
Query Match: 64.4% Indels: 7
DB: 1 Gaps: 1

US-10-664-025-3903 (1-153) x AI554902 (1-480)

Qy 20 LeuHisValLysIleSerCysSerLeuAspTyrLeuMetValSerValIleProValAla 39
480 CTCATGTGAAATAAGTTGCTCTCTGGAATGCTGATGCTCAGTTATCCAGTTGCA 421
Qy 40 GluSerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla 59
420 GAAAGCAGAAATCTGTATATATTTGCGATGAATTACATCTGGGAATGGGCTGCCCTGCA 361
Qy 60 AsnArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArg 79
360 AATCGATACATACATATGATATGATATATATATCTTGTTCGTGATTTGGCATCAGG 301
Qy 80 ThrArgValValSerGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArg 99
300 ACAAGGGTAGTTTCTGAGGAAACTCTCTCTTTTCAAACCGAGCTGTACTTTACCCCAAGG 241
Qy 100 AsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArg----- 116
240 AATATAGATCATGACCTCAGGAAATCCATTTGGAGTGTTCACCTCTAGTGTGTGATGC 181
Qy 117 -----LysSerValTyr 120
180 AGAATCATGGGAGAGACTATCTGG 157

RESULT 11

BG081002

LOCUS

DEFINITION

H3059H12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone

H3059H12 5', mRNA sequence.

ACCESSION

BG081002

VERSION

BG081002.2 GI:40015217

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 598)

AUTHORS

Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H. III, Becker,K.G. and Ko,M.S.H.

TITLE

Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

PUBMED

10922068

COMMENT

On Jan 26, 2001 this sequence version replaced gi:12563570.

Other_ESTs: H3059H12-3

Contact: George J. Kargul

Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3059 row: H column: 12
Seq primer: -21M13 Reverse
High quality sequence stop: 598
POLYA=No.

FEATURES	Source
Location/Qualifiers	
1..598	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6J"	
/db_xref="niaEST:H3059H12-5"	
/db_xref="taxon:10090"	
/clone="H3059H12"	
/sex="Clones arrayed from a variety of cDNA libraries"	
/dev_stage="Clones arrayed from a variety of cDNA libraries"	
/lab_host="DH10B"	
/clone_lib="NIA Mouse 15K cDNA Clone Set"	
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation of the X chromosome. 1998. Hum Mol Genet 7: 1967-1978."	

ORIGIN

Alignment Scores:	
Pred. No.:	2,28e-42
Score:	431.50
Length:	598
Matches:	85
Conservative:	25
Mismatches:	41
Indels:	1
Gaps:	1
DB:	2

US-10-664-025-3903 (1-153) x BG081002 (1-598)

	Qy	1	MetAlaLeuGluValLeuMetLeuLeuAlaValLeulleTrpThrGlyAlaGluAsnLeu	20
	Dd	41	AATGGCTTTAGAGGTCTTGCGTATTCTTCCTCTCTTGTTTTGGACATTGTGCTTGGACAATT	100
	Qy	21	HISValLysIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu	40
	Dd	101	GACGTGGATGAAGTGTCTTCAGACTGGAGTAGCGGTCA GTGTAGTGGCTTTCCAG	160
	Qy	41	Ser--ArgAsnLeuTyrlpheAlaaspGluLeuHisLeuGlyMetGlyCysProAla	59
	Dd	161	AACAAAAAGAATAACCATAATTTTGCTCATGAATTATATCTGGACAGAACTGCCGTGTA	220
	Qy	60	AsnA r g I l e H i s T h r T y r V a l T y C l u p h e l l e t y r L e u V a l A r g A s p C y s G l y I l e A r g	79
	Dd	221	ACTCAAATACAGCGACCAATATGATTTCATATACCCTGTTCTCTACTGTGGGCATCAGG	280
	Qy	80	ThrArgValValserGluGluThrLeuLeuPheGlnThrGluLeuTyrlPheThrProArg	99
	Dd	281	ACAAAGGTATATCTCAAAAGAAATCGTCNGTTTTCAGAACCCAGGATGTA CTTCGTCCAAGG	340


```
US-10-664-025-3903 (1-153) x CO810251 (1-862)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTyrThrGlyAlaGluAsnLeu 20
Db 8 ATGGCTTTAGAGTCTGCTGTTATCTGCTGCTGTTGTTGACTGTGCTGGGACATT 67
QY 21 HisValIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 68 GAGCTGGAGTAAAGTTGTTCTCAGGACTGGATCGGTCAGTGTAGTGGTTCCTCCAG 127
QY 41 Ser---ArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAla 59
Db 128 AACAAAAGAAATCCATATATTTTGTGATGAATTATATCTGGGACAGAACTGCCGTGTA 187
QY 60 AsnArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArg 79
Db 188 ACTCAAAATACAGCATCAATGATTTCAATATACCTGTTCTCAGTGTGGCATCAGG 247
QY 80 ThrArgValIleSerGluGluThrLeuLeuPheGlnThrGluLeuTyrPheThrProArg 99
Db 248 ACAAGGTTATCTCAATGAATCGTCTGTTTGAACCGAGATGTACTTTCGTCCAAG 307
QY 100 AsnIleAspHisAspProGlnGluIleHisLeuGluCysSerThrSerArgIleSerVal 119
Db 308 AATTATTTGTTGGAACCTTCAGATAGTCCCTTTGCGAGTGTCTGCCCTCTCGGAAGTCTGTG 367
QY 120 TrpLeuThrProValSerThrGluAsnGluIleIleLeuLeuAspProSerProPheIleAla 139
Db 368 TGGCTTATCCCACTATCTACTGAGGAGGATCCAAACCTGTTAAAGAGTCCCTTTATGACT 427
QY 140 AspPheGlnThrThrAlaGluGluLeuGlyLeuLeu 151
Db 428 GACTTTGAGCTACACAGAGATGGGGTTATTA 463

RESULT 14
A1138881/c
LOCUS
DEFINITION
q99c04.x1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1737606
3', mRNA sequence.
ACCESSION
A1138881
VERSION
A1138881.1 GI:3644853
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 331)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 395 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
FEATURES
source
1..331
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1737606"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares_testis_NHT"
```

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

```
Alignment Scores:
Pred. No.: 2,09e-42 Length: 331
Score: 429.00 Matches: 82
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 53.7% Indels: 0
DB: 1 Gaps: 0
US-10-664-025-3903 (1-153) x A1138881 (1-331)
QY 1 MetAlaLeuGluValLeuMetLeuLeuAlaValLeuLeuTyrThrGlyAlaGluAsnLeu 20
Db 331 ATGGCGTTAGAAAGTCTTGATGCTCCTCGCTGCTTGTGATTGGACCGGTCTGAGAACCCTC 272
QY 21 HisValIleIleSerCysSerLeuAspTrpLeuMetValSerValIleProValAlaGlu 40
Db 271 CATGTGAAATAAGTTGCTCTCTGAGCTGGTTGATGGTCTCAGTTATCCAGTTGCAGAA 212
QY 41 SerArgAsnLeuTyrIlePheAlaAspGluLeuHisLeuGlyMetGlyCysProAlaAsn 60
Db 211 AGCAGAATCTGTATATATTTGCGGATGAATTACATCTGGGAATGGGCTGCCCTCAAT 152
QY 61 ArgIleHisThrTyrValTyrGluPheIleTyrLeuValArgAspCysGlyIleArgThr 80
Db 151 CGGATACATACATATGATATATGAGTTATATATCTTGTGTTGATTTGGTCATCAGGACA 92
QY 81 ArgVal 82
Db 91 AGGGTA 86
RESULT 15
BB557840 642 bp mRNA linear EST 26-OCT-2001
LOCUS
BB557840 RIKEN full-length enriched, 2 days pregnant adult female
ovary Mus musculus cDNA clone E330031A07 3', mRNA sequence.
ACCESSION
BB557840
VERSION
BB557840.2 GI:16448837
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 642)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
JOURNAL
On Aug 1, 2000 this sequence version replaced gi:9644206.
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
```


This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:10:22 ; Search time 105.3 Seconds
(without alignments)
638.414 Million cell updates/sec

Title: US-10-664-025-3903
Perfect score: 799
Sequence: 1 MALEVLMLLAVLWTCAGNL.....PSPFYADFOTTAELGLLIF 153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	789	98.7	158	4	AAG89140 Human sec
2	789	98.7	158	5	ABG97353 Human CGD
3	789	98.7	158	5	ABG31324 Human GSS
4	789	98.7	158	6	ABR47931 Human sec
5	789	98.7	158	6	ABR00175 Human gen
6	789	98.7	158	7	ADB91675 Human sec
7	789	98.7	158	7	ADC74338 Human sec
8	789	98.7	159	3	AAB38337 Human sec
9	785	98.2	158	7	ADB65520 Human pro
10	775	97.0	148	6	ABP76237 Human GEN
11	752	94.1	143	6	ABP76237 Human GEN
12	429	53.7	106	4	AAG89145 Human sec
13	225.5	28.2	212	7	ADB28663 Human NOV
14	225.5	28.2	212	8	ADM93408 Human NOV
15	223.5	28.0	212	2	AAW73630 Human sec
16	223.5	28.0	212	3	AAY94302 Human cor
17	223.5	28.0	212	4	AAm93517 Human pol
18	223.5	28.0	212	5	ABP62038 Human sec
19	223.5	28.0	212	5	ABP61861 Human pol
20	223.5	28.0	212	8	ADC24649 Protein e
21	223.5	28.0	212	8	ADJ31210 Human pro
22	223.5	28.0	213	7	AAY12939 Amino aci
23	183.5	23.0	139	7	ABR83176 PLAC1 int
24	161	20.2	163	7	ADB28665 Human NOV

25	161	20.2	163	8	ADM93410	Human NOV
26	98.5	12.3	930	4	ABG12048	Novel hum
27	96.5	12.1	139	7	ABR83177	PLAC1 int
28	94	11.8	59	4	ABB21608	Protein #
29	94	11.8	59	4	ABG51070	Human liv
30	88.5	11.1	234	4	ABG27342	Novel hum
31	85.5	10.7	1062	8	ADR09311	Human pro
32	85.5	10.7	3648	8	ADP30076	Human sec
33	85.5	10.7	3776	8	ADP30075	Human sec
34	85.5	10.7	5373	4	AAU14603	Novel bon
35	85.5	10.7	5373	7	ADJ68935	Human hea
36	85.5	10.7	5430	9	ADZ85102	Full-leng
37	85.5	10.7	5447	4	AAU14697	Novel bon
38	84	10.5	1241	6	ABU18092	Protein e
39	79.5	9.9	431	4	AAG85042	Shrimp wh
40	77	9.6	358	5	ABB82478	Rat atear
41	77	9.6	358	7	ADE61071	Rat Prote
42	77	9.6	358	8	ADI79900	Human E1-
43	76	9.5	158	8	ADK46733	Streptoco
44	76	9.5	165	8	ADR95319	Novel S.
45	76	9.5	165	9	AEA59189	Streptoco

ALIGNMENTS

RESULT 1
AAG89140
ID AAG89140 standard; protein; 158 AA.
XX
AC AAG89140;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 260.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX (GEST) GENSET.
Dumas Milne Edwards J, Bougueleret L, Jobert S;
WPI: 2001-367870/38.
DR N-PSDB; AAH64743.
Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases.
Claim 21; Page 804; 921pp; English.
The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their

CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET polypeptide of
CC the invention

XX Sequence 158 AA;

SQ Query Match 98.7%; Score 789; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDLWLVSVIPVAESRNLVYFADELHLMGCGPAN 60
DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDLWLVSVIPVAESRNLVYFADELHLMGCGPAN 60
QY 61 RIHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTRKSVW 120
DB 61 RIHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTRKSVW 120
QY 121 LTPVSTENEIKLDPSPFIADFTQTAEELGLL 151
DB 121 LTPVSTENEIKLDPSPFIADFTQTAEELGLL 151

RESULT 2
ABG97353
ID ABG97353 standard; protein; 158 AA.

XX AC ABG97353;

DT 16-DEC-2002 (first entry)

XX DE Human CGDD4, INCYTE 5284076CD1.

XX Human; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.

XX OS Homo sapiens.

XX PN WO200272830-A2.

XX PD 19-SEP-2002.

XX PF 08-FEB-2002; 2002WO-US003715.

XX PR 09-FEB-2001; 2001US-0268111P.

XX PR 23-FEB-2001; 2001US-0271175P.

XX PR 08-MAR-2001; 2001US-0274503P.

XX PR 09-MAR-2001; 2001US-0274552P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;

PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;

PI Lu DAM, Richardson TW, Tran UK, Khare R, Wallia NK;

XX WPI; 2002-723356/78.

XX DR N-PSDB; ABS78646.

PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.

XX Claim 1; Page 146; 181pp; English.

XX The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence
CC represents a CGDD protein

XX Sequence 158 AA;

Query Match 98.7%; Score 789; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDLWLVSVIPVAESRNLVYFADELHLMGCGPAN 60
DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDLWLVSVIPVAESRNLVYFADELHLMGCGPAN 60

QY 61 RIHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTRKSVW 120
DB 61 RIHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTRKSVW 120

QY 121 LTPVSTENEIKLDPSPFIADFTQTAEELGLL 151

DB 121 LTPVSTENEIKLDPSPFIADFTQTAEELGLL 151

RESULT 3

ABG31324
ID ABG31324 standard; protein; 158 AA.

XX AC ABG31324;

XX DT 05-NOV-2002 (first entry)

XX DE Human GSSP3 polypeptide.

XX Human; GSSP3; circulating blood glucose level; insulin sensitivity;
KW body mass; serum glucose regulation; body weight loss; obesity;
KW metabolic-related disorder; impaired glucose tolerance; stroke;
KW insulin resistance; hyperlipidaemia; atherosclerosis; heart disease;
KW hypertension; syndrome C; type I diabetes; type II diabetes;
KW microangiopathy; retinopathy; neuropathy; nephropathy; ocular lesion;
KW polycystic ovarian syndrome; acanthosis nigricans; leprechaunism;

KW lipoatrophy; physical performance; exercise; dyslexia; schizophrenia;
KW attention-deficit disorder; hyperactivity disorder; psychiatric disorder.
XX
XX Homo sapiens.

XX Key Location/Qualifiers
XX Peptide 1. .17
XX FT /label= signal_peptide
XX Protein 18 .158
XX FT /label= Mature_GSSP3_polypeptide

XX WO200260466-A2.

XX
XX
XX

XX 08-AUG-2002.

XX 01-FEB-2002; 2002WO-IB001333.

XX 02-FEB-2001; 2001US-0266156P.

XX (GEST) GENSET.

XX Salter-Cid L, Ebbets-Reed D, Bour BA, Chicca J, Yen-Potin F;

XX Bihain B;

XX WPI; 2002-608487/65.

XX N-PSDB; ABK90053.

XX Reducing circulating glucose levels or increasing insulin sensitivity,
XX useful for reducing body mass or preventing body weight gain, comprises
XX administering composition comprising GSSP3 polypeptide.

XX Claim 1; Page 96-97; 97pp; English.

XX The present invention relates to the isolation of human GSSP3
XX polypeptide, and polynucleotide sequences that encode it. The GSSP3
XX polypeptide reduces circulating blood glucose levels, increases insulin
XX sensitivity, and/or reduces body mass. The GSSP3 polypeptide and
XX polynucleotide sequences are useful in serum glucose regulation, fatty
XX acid metabolism, body weight loss, and prevention of body weight gain.
XX Compositions comprising GSSP3 polypeptides are useful for controlling
XX blood glucose levels, for treating metabolic-related diseases or
XX disorders (e.g. obesity, impaired glucose tolerance, insulin resistance,
XX hyperlipidaemias, atherosclerosis, heart disease, hypertension, stroke,
XX syndrome C, type I or II diabetes, diabetes related complications,
XX microangiopathy, retinopathy, neuropathy, nephropathy, polycystic ovarian
XX syndrome, ocular lesions, microangiopathic lesions, or syndromes such as
XX acanthosis nigricans, leprechaunism and lipodystrophy). The polypeptides
XX are also useful to improve physical performance during work or exercise,
XX and to treat dyslexia, attention-deficit disorder, attention-
XX deficit/hyperactivity disorder, and psychiatric disorders such as
XX schizophrenia. The present sequence represents human GSSP3 polypeptide

XX Sequence 158 AA;

Query Match 98.7%; Score 789; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHGMGCPAN 60

DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYEFYILVRDCGIRTVVSEETLLFQTELYFTPRNIDHPDQEIHLECSRSKSW 120

DB 61 RIHTVYEFYILVRDCGIRTVVSEETLLFQTELYFTPRNIDHPDQEIHLECSRSKSW 120

QY 121 LTPVSTENEIKLDPSPFIADPQTABELGL 151

DB 121 LTPVSTENEIKLDPSPFIADPQTABELGL 151

RESULT 4
ABR47931

ID
XX AC

ABR47931 standard; protein; 158 AA.

XX ABR47931;

XX 12-JUN-2003 (first entry)

XX Human secreted protein, SEQ ID 822.

XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
XX vulnery; antiinflammatory; nootropic; neuroprotective;
XX antiparkinsonian; gene therapy; human; cardiovascular disorder.

XX Homo sapiens.

XX WO200295010-A2.

XX 28-NOV-2002.

XX 19-MAR-2002; 2002WO-US009785.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-129429/12.

XX Novel human secreted proteins, useful for detecting, preventing,
XX diagnosing, prognosticating, treating and/or ameliorating cardiovascular
XX disorders such as arrhythmia.

XX Claim 13; SEQ ID NO 822; 1881pp; English.

XX The present invention relates to novel human secreted proteins (ABR47633-
XX ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
XX and their coding sequences are useful for the preparation of a diagnostic
XX or pharmaceutical composition for diagnosing or treating a cardiovascular
XX disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
XX arteriosclerosis and myocardial ischaemia), neural disorders, immune
XX system disorders, muscular disorders, reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, renal disorders,
XX proliferative disorders and/or cancerous diseases and conditions, for
XX wound healing and epithelial cell proliferation, to treat inflammation or
XX infection, for treating thrombosis and arteriosclerosis, for treating or
XX preventing neural damage which occurs in neuronal disorders or
XX neurodegenerative conditions such as Alzheimer's disease and Parkinson's
XX disease, to enhance bone and periodontal regeneration and aid in tissue
XX transplants or bone grafts, to prevent skin aging or hair loss, to
XX stimulate growth and differentiation of haematopoietic cells and bone
XX marrow cells when used in combination with other cytokines, to maintain
XX organs before transplantation or for supporting cell culture of primary
XX tissues, to increase or decrease differentiation or proliferation of
XX embryonic stem cells, or to modulate mammalian characteristics or
XX metabolism. Note: The sequence data for this patent was published in
XX electronic format and is available from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 158 AA;

Query Match 98.7%; Score 789; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHGMGCPAN 60

DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYEFYILVRDCGIRTVVSEETLLFQTELYFTPRNIDHPDQEIHLECSRSKSW 120

DB 61 RIHTVYEFYILVRDCGIRTVVSEETLLFQTELYFTPRNIDHPDQEIHLECSRSKSW 120

```
QY      121 LTPVSTENEIKLDPSPFFIADFQTAAEELGLL 151
      |||||
Db      121 LTPVSTENEIKLDPSPFFIADFQTAAEELGLL 151

RESULT 5
ID ABR00175
XX ABR00175 standard; protein; 158 AA.
AC ABR00175;
XX
DT 03-APR-2003 (first entry)
XX
DE Human gene 165 encoded secreted protein HTELS08, SEQ ID NO:464.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulneryary; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200276488-A1.
XX
PD 03-OCT-2002.
XX
PF 19-MAR-2002; 2002WO-US008276.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-029900/02.
DR N-PSDB; ABZ71354.
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers.
XX
PS Claim 13; Page 1048; 1216pp; English.
XX
CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein of the
CC invention.
XX
SQ Sequence 158 AA;

Query Match      98.7%; Score 789; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWTGAENLHVKJSCSLDWLWVSVIPVAESRNLVIFADELHLGMCSPAN 60
      |||||
Db      1 MALEVLMLLAVLIWTGAENLHVKJSCSLDWLWVSVIPVAESRNLVIFADELHLGMCSPAN 60

QY      61 RIHTYVVEFIYLVRCDCGIRTVVSEETLLFQTELYFTPRNIDHDHPQEIHLECSTSRKSVW 120
      |||||
Db      61 RIHTYVVEFIYLVRCDCGIRTVVSEETLLFQTELYFTPRNIDHDHPQEIHLECSTSRKSVW 120

QY      121 LTPVSTENEIKLDPSPFFIADFQTAAEELGLL 151
      |||||
Db      121 LTPVSTENEIKLDPSPFFIADFQTAAEELGLL 151

RESULT 6
ADB91675
ID ADB91675 standard; protein; 158 AA.
XX
AC ADB91675;
XX
DT 04-DEC-2003 (first entry)
XX
XX Human secreted protein #SEQ ID 621.
XX
KW Secreted protein; gene therapy; antidiabetic; diabetes; human.
XX
OS Homo sapiens.
XX
PN WO2003004622-A2.
XX
PD 16-JAN-2003.
XX
PF 19-MAR-2002; 2002WO-US008124.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-229407/22.
XX
PT Nucleic acid encoding a human secreted protein is useful in diagnosing or
PT treating diabetes or conditions related to diabetes.
XX
PS Claim 3; SEQ ID NO 621; 1537pp; English.
XX
CC The invention relates to isolated nucleic acid molecules ADB91065-
CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
CC ADB91834. Also disclosed is a recombinant vector comprising a
CC polynucleotide of the invention, and a recombinant host cell comprising
CC the recombinant vector. The polypeptide of the invention is useful in
CC identifying a binding partner by contacting the polypeptide with a
CC binding partner, and determining whether the binding partner increases or
CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
CC antibody or its fragment, agonist or antagonist are useful for preparing
CC a pharmaceutical composition for diagnosing or treating diabetes or
CC conditions related to diabetes. The present sequence is that of the human
CC immunoglobulin Fc portion used to generate fusion proteins, increasing
CC the stability of the fused protein as compared to the secreted protein
CC only. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 158 AA;

Query Match      98.7%; Score 789; DB 7; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.3e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHLMGCPAN 60
 |||||
 Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHLMGCPAN 60
 |||||
 QY 61 RIHTYVEFIYLRDVGIRTRVSVSEETLLFQTELYFTPRNIDHPQEIHLHCSTSRKSVW 120
 |||||
 Db 61 RIHTYVEFIYLRDVGIRTRVSVSEETLLFQTELYFTPRNIDHPQEIHLHCSTSRKSVW 120
 |||||
 QY 121 LTPVSTENEIKLDPSPFIADFQTTAEELGLL 151
 |||||
 Db 121 LTPVSTENEIKLDPSPFIADFQTTAEELGLL 151
 |||||
 RESULT 7
 ADC74338
 ID ADC74338 standard; protein; 158 AA.
 XX
 AC ADC74338;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human secreted protein - SEQ ID 971.
 XX
 KW antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
 KW antidiabetic; immunosuppressive; dermatologic; nephrotropic;
 KW antiparkinsonian; neuroprotective; nontropic; antibacterial; virucide;
 KW fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;
 KW haemopoietic; haematologic; anaemia; autoimmune disorder;
 KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
 KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
 KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
 KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003038063-A2.
 XX
 PD 08-MAY-2003.
 XX
 PF 19-MAR-2002; 2002WO-US008277.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2003-430516/40.
 DR N-PSDB; ADC73723.
 XX
 PT New human secreted polypeptide for diagnosing, preventing or treating
 PT hematopoietic or hematologic disorders (e.g. anemia), autoimmune
 PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
 PT atherosclerosis).
 XX
 PS Claim 16; SEQ ID NO 971; 2272pp; English.
 XX
 CC The invention relates to a novel human secreted polypeptide comprising a
 CC defined sequence given in the specification. The polypeptide, nucleic
 CC acid molecule, antibody, agonist or antagonist of the invention may be
 CC useful for preparing a composition for diagnosing or treating a
 CC haemopoietic or haematologic disorder such as anaemia, autoimmune
 CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
 CC diabetes, systemic lupus erythematosus or glomerulonephritis,
 CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
 CC disease, wounds and hyperproliferative disorders including
 CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
 CC parasitic infections. The polypeptide may also be used during gene
 CC therapy procedures and for identifying a binding partner by contacting
 CC the polypeptide with a binding partner and determining whether the

CC binding partner increases or decreases the activity of the polypeptide.
 CC The current sequence is that of the human secreted protein of the
 CC invention.
 XX
 SQ Sequence 158 AA;
 Query Match 98.7%; Score 789; DB 7; Length 158;
 Best Local Similarity 100.0%; Pred. No. 3.3e-81;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHLMGCPAN 60
 |||||
 Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHLMGCPAN 60
 |||||
 QY 61 RIHTYVEFIYLRDVGIRTRVSVSEETLLFQTELYFTPRNIDHPQEIHLHCSTSRKSVW 120
 |||||
 Db 61 RIHTYVEFIYLRDVGIRTRVSVSEETLLFQTELYFTPRNIDHPQEIHLHCSTSRKSVW 120
 |||||
 QY 121 LTPVSTENEIKLDPSPFIADFQTTAEELGLL 151
 |||||
 Db 121 LTPVSTENEIKLDPSPFIADFQTTAEELGLL 151
 |||||
 RESULT 8
 AAB38337
 ID AAB38337 standard; protein; 159 AA.
 XX
 AC AAB38337;
 XX
 DT 31-JAN-2001 (first entry)
 XX
 DE Human secreted protein encoded by gene 17 clone HTLS08.
 XX
 KW Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
 KW cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;
 KW nontropic; antibacterial; virucide; fungicide; ophthalmological; human;
 KW vulnerary; gene therapy; infection; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO2000061623-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 06-APR-2000; 2000WO-US008979.
 XX
 PR 09-APR-1999; 99US-0128693P.
 PR 26-APR-1999; 99US-0130991P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ni J, Komatsoulis GA, Rosen CA, Soppet DR, Shi Y;
 PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Moore PA, Birse CE;
 PI Young PE;
 XX
 DR WPI: 2000-647418/62.
 XX
 PT New nucleic acid molecules encoding 62 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 XX
 PS Claim 11; Page 608-609; 716pp; English.
 XX
 CC Sequences AAB38321-B38396 represent the amino acid sequences of 62 human
 CC secreted proteins encoded by the genes AAC69512-C69587. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.
 CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms of
 CC the breast or liver; (c) cardiovascular disorders e.g. cardiac arrest;
 CC (d) cerebrovascular disorders e.g. cerebral ischemia; (e) angiogenesis;

CC (f) nervous system disorders e.g. Alzheimer's disease; (g) infections
CC caused by bacteria, viruses and fungi; and (h) ocular disorders e.g.
CC corneal infection. The polypeptides can also be used to aid wound healing
CC and epithelial cell proliferation, to prevent skin aging due to sunburn, or
CC to maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis
XX
SQ Sequence 159 AA;

Query Match 98.7%; Score 789; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.4e-81;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMCCPAN 60
DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMCCPAN 60
QY 61 RIHTVYVEFIYLRDCGIRTRVSEETLLFQTELYFTPRNIDHDPQEIHLCESTSRKSVW 120
DB 61 RIHTVYVEFIYLRDCGIRTRVSEETLLFQTELYFTPRNIDHDPQEIHLCESTSRKSVW 120
QY 121 LTPVSTENEIKLDPSFPFIADFQTAAEELGLL 151
DB 121 LTPVSTENEIKLDPSFPFIADFQTAAEELGLL 151

RESULT 9
ADB65520
ID ADB65520 standard; protein; 158 AA.
XX
AC ADB65520;
DT 04-DEC-2003 (first entry)
XX
DE Human protein encoded by clone TEST120282420.
XX
KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX
OS Homo sapiens.
XX
PN EP1308459-A2.
XX
PD 07-MAY-2003.
XX
PF 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehli S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
WPI; 2003-450961/43.
DR N-PSDB; ADB63550.
XX
PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page; 222pp; English.
XX

CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or

CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesising the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 158 AA;

Query Match 98.2%; Score 785; DB 7; Length 158;
Best Local Similarity 99.3%; Pred. No. 9.5e-81;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMCCPAN 60
DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMCCPAN 60
QY 61 RIHTVYVEFIYLRDCGIRTRVSEETLLFQTELYFTPRNIDHDPQEIHLCESTSRKSVW 120
DB 61 RIHTVYVEFIYLRDCGIRTRVSEETLLFQTELYFTPRNIDHDPQEIHLCESTSRKSVW 120
QY 121 LTPVSTENEIKLDPSFPFIADFQTAAEELGLL 151
DB 121 LTPVSTENEIKLDPSFPFIADFQTAAEELGLL 151

RESULT 10
ABP76238
ID ABP76238 standard; protein; 148 AA.
XX
AC ABP76238;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET protein SEQ ID 788.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2001; 2001WO-IB000914.
XX
PR 18-APR-2001; 2001WO-IB000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX
WPI; 2003-075548/07.
XX
PT New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the

PT toxicity.
XX Claim 14; Page 680; 735pp; English.
XX
CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for
CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity
XX
XX Sequence 148 AA;
Query Match 97.0%; Score 775; DB 6; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.2e-79;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLVYFADSLHGMGCPAN 60
DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLVYFADSLHGMGCPAN 60
QY 61 RIHTYVVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCESTRKSVW 120
DB 61 RIHTYVVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCESTRKSVW 120
QY 121 LTPVSTENEIKLDPSPFIADFTABEL 148
DB 121 LTPVSTENEIKLDPSPFIADFTABEL 148
RESULT 11
ABP76237
ID ABP76237 standard; protein; 143 AA.
XX
AC ABP76237;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human GENSET protein SEQ ID 787.
XX
KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
KW inflammatory disease; immune disorder; neuromuscular; toxicity;
KW central nervous system; cardiovascular; gastrointestinal.
XX
OS Homo sapiens.
XX
PN WO200283898-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2001; 2001WO-IB000914.
XX
PR 18-APR-2001; 2001WO-IB000914.
XX
PA (GEST) GENSET.
XX
PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
XX
XX WPI; 2003-075548/07.
XX
XX New GENSET polynucleotides and polypeptides, useful for treating heavy
PT metal toxicity, cancer, inflammatory diseases, immune disorders, and the
PT neuromuscular, CNS, cardiovascular or gastrointestinal effects of the
PT toxicity.
XX
XX Claim 14; Page 680; 735pp; English.
XX
CC The present invention relates to novel GENSET polynucleotides (ABZ36404-
CC ABZ36911) encoding polypeptides (ABP75963-ABP76368). The polynucleotides
CC and polypeptides are useful in screening and diagnostic assays for

CC abnormal GENSET expression and/or biological activity. They are also
CC useful for screening of compounds for treating or preventing GENSET-
CC related disorders, such as heavy metal toxicity, cancer, inflammatory
CC diseases, immune disorders, and the neuromuscular, central nervous system
CC (CNS), cardiovascular or gastrointestinal effects of the toxicity
XX
XX Sequence 143 AA;
Query Match 94.1%; Score 752; DB 6; Length 143;
Best Local Similarity 100.0%; Pred. No. 4.7e-77;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLVYFADSLHGMGCPAN 60
DB 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLVYFADSLHGMGCPAN 60
QY 61 RIHTYVVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCESTRKSVW 120
DB 61 RIHTYVVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCESTRKSVW 120
QY 121 LTPVSTENEIKLDPSPFIADFTQ 143
DB 121 LTPVSTENEIKLDPSPFIADFTQ 143
RESULT 12
AAG89145
ID AAG89145 standard; protein; 106 AA.
XX
AC AAG89145;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 265.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB001938.
XX
PR 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
XX WPI; 2001-367870/38.
DR N-PSDB; AAH64748.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
PS Claim 21; Page 806; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patient's own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA

KW acquired immunodeficiency syndrome; AIDS; Crohn's disease;
XX infectious disease; anorexia; immune disorder.
OS Homo sapiens.

XX US2004067882-A1.
XX 08-APR-2004.

XX 05-NOV-2002; 2002US-00287971.
XX 22-OCT-2001; 2001US-00035568.

XX 05-NOV-2001; 2001US-0338626P.
XX 06-NOV-2001; 2001US-0333072P.

XX 09-NOV-2001; 2001US-0345398P.
XX 09-NOV-2001; 2001US-0348283P.

XX 13-NOV-2001; 2001US-0335610P.
XX 21-NOV-2001; 2001US-0332152P.

XX 28-NOV-2001; 2001US-0333912P.
XX 29-NOV-2001; 2001US-00997425.

XX 29-NOV-2001; 2001US-0334300P.
XX 04-DEC-2001; 2001US-0336576P.

XX 05-FEB-2002; 2002US-0354807P.
XX 15-MAY-2002; 2002US-0380968P.

XX 16-MAY-2002; 2002US-0381043P.
XX 02-JUL-2002; 2002US-0393148P.

XX 02-JUL-2002; 2002US-0393262P.
XX 08-AUG-2002; 2002US-0401479P.

XX 06-AUG-2002; 2002US-0401626P.
XX 07-AUG-2002; 2002US-0401593P.

XX 07-AUG-2002; 2002US-0401695P.
XX 26-AUG-2002; 2002US-0406181P.

(ALSO/) ALSOBROOK J P.

(ALVA/) ALVAREZ E.

(ANDE/) ANDERSON D W.

(BARO/) BARON M.

(BOLD/) BOLDOG F L.

(BURG/) BURGESS C E.

(CASW/) CASMAN S J.

(CHAP/) CHAPOVAL A.

(DHAN/) DHANABAL M.

(EDIN/) EDINGER S R.

(EISE/) EISEN A.

(ELLE/) ELLERMAN K.

(ETTE/) ETENBERG S.

(GANG/) GANGOLLI E A.

(GERL/) GERLACH V.

(GORM/) GORMAN L.

(GROS/) GROSSE W M.

(GUOX/) GUO X.

(HACK/) HACKETT C.

(JIWV/) JI W.

(KEKU/) KEKUDA R.

(KHRA/) KHRAMTSOV N V.

(LEPL/) LEPLEY D M.

(LILL/) LI L.

(MACD/) MACDOUGALL J R.

(MALY/) MALYANKAR U M.

(MAZU/) MAZUR A.

(MCQU/) MCQUEENEY K.

(MEZE/) MEZES P S.

(MILL/) MILLER C E.

(MILL/) MILLER I.

(MISH/) MISHRA V.

(PADI/) PADIGARU M.

(PATT/) PATTURAJAN M.

(PENA/) PENA C E A.

(PEYN/) PEYMAN J A.

(RASI/) RASTELLI L.

(RIEG/) RIEGER D K.

PA (SHIM/) SHIMKETS R A.
PA (SMIT/) SMITHSON G.
PA (SPAD/) SPADERNA S K.
PA (STAR/) STARLING G.
PA (SPYT/) SPYTEK K A.
PA (STON/) STONE D J.
PA (TCHE/) TCHERNEV V T.
PA (TWOI/) TWOLOW N.
PA (VERN/) VERNET C A M.
PA (ZERR/) ZERHUSEN B D.
PA (VOSS/) VOSS E Z.
PA (ZHON/) ZHONG M.

PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI Burgess CE, Casman SU, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L;
PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
PI Lepley DM, Li L, MacDougall JR, Malyankar UM, Mazur A, McQueeney K;
PI Mezes PS, Miller CE, Millet I, Mishra V, Padigar M, Patturajan M;
PI Pena CE, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME;
PI Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G;
PI Spytek KA, Stone DJ, Tchernev VT, TwoLOW N, Vernet CM;
PI Zerhusen BD, Voss EZ, Zhong M;

WPI: 2004-355303/33.
N-PSDB; ADM93407.

Novel isolated NOVX polypeptide useful treating or preventing disorders
or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.

Claim 2; SEQ ID NO 40; 330pp; English.

The invention relates to human NOVX polypeptides and the polynucleotides
encoding them. The NOVX polypeptides and polynucleotides are useful for
determining the presence of or predisposition to a disease associated
with altered levels of the sequences in a mammalian subject, and for
treating or preventing a pathology associated with NOVX. The
polypeptides, polynucleotides and antibodies that bind immunospecifically
to the polypeptides are useful for treating or preventing disorders or
syndromes such as congenital heart defects, cardiomyopathy,
atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
anorexia and immune disorders. This sequence represents a human NOVX
polypeptide of the invention. Note: The sequence data for this patent is
also available from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 212 AA;

Query Match 28.2%; Score 225.5; DB 8; Length 212;
Best Local Similarity 33.3%; Pred. No. 6.1e-17;
Matches 40; Conservative 33; Mismatches 42; Indels 5; Gaps 2;

QY 5 VLMLAVLIWTGAENLHVKISCDLWVSVIPVAESRLNYIFADELHLMGCPANRIHT 64
Db 9 LMLILTSALSGSQSPMTVLCSIDWVTVHPFNLNDVCFVHFLHGLGCPNHHVQ 68
QY 65 YVVEFIYLRDCGIRTVVSEETLLFQTELYTPRNIDHPDE--IHLECSYRSKSVLIT 122
Db 69 HAYQTYRVTGIRAKAVSQDMVIYSTEIHYSSKG---TPSKFVIPVSCAAPQKSPWLT 125

RESULT 15

AAW73630

ID AAW73630 standard; protein; 212 AA.

XX AC AAW73630;

XX DT 23-MAR-1999 (first entry)

XX

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:14:54 ; Search time 36 Seconds
(without alignments)
351.372 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLIWTGAENL.....PSPFIADFTTAEELGLLIF 153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	799	100.0	153	2	US-09-621-976-3903
2	799	98.7	158	2	US-09-621-976-5222
3	785	98.2	158	2	US-10-104-047-3674
4	429	53.7	106	2	US-09-621-976-5223
5	223.5	28.0	216	2	US-09-949-016-11232
6	211	26.4	56	2	US-09-621-976-6447
7	77	9.6	358	1	US-08-558-823-19
8	77	9.6	358	2	US-09-723-806A-8
9	76	9.5	158	2	US-09-583-110-3248
10	76	9.5	165	2	US-09-107-433-3954
11	76	9.5	323	2	US-09-489-039A-10162
12	75.5	9.4	385	2	US-09-328-352-7887
13	75.5	9.4	424	1	US-08-484-993B-18
14	75.5	9.4	424	1	US-08-484-158B-18
15	75.5	9.4	424	1	US-08-484-596A-18
16	75.5	9.4	424	1	US-08-480-150A-18
17	75.5	9.4	424	2	US-08-458-731-18
18	75.5	9.4	424	2	US-08-149-223A-18
19	74.5	9.3	424	1	US-08-453-472-8
20	74.5	9.3	424	1	US-08-038-948-2
21	74.5	9.3	424	1	US-08-038-948-5
22	74.5	9.3	424	1	US-08-453-952-8
23	74.5	9.3	424	1	US-08-862-903-8
24	72.5	9.1	156	2	US-09-270-767-36988
25	72.5	9.1	156	2	US-09-270-767-52205
26	71.5	8.9	649	2	US-10-104-047-3492
27	71.5	8.9	1020	1	US-08-070-301-3

28 71 8.9 288 2 US-09-949-016-10862 Sequence 10862, A
29 71 8.9 424 1 US-08-453-472-7 Sequence 7, Appli
30 71 8.9 424 1 US-08-038-948-4 Sequence 4, Appli
31 71 8.9 424 1 US-08-453-952-7 Sequence 7, Appli
32 71 8.9 424 1 US-08-862-903-7 Sequence 7, Appli
33 71 8.9 424 1 US-08-484-158B-61 Sequence 61, Appli
34 70.5 8.8 519 2 US-09-312-183A-3 Sequence 3, Appli
35 70.5 8.8 528 2 US-09-489-039A-13077 Sequence 13077, A
36 70.5 8.8 547 2 US-09-312-183A-2 Sequence 2, Appli
37 70.5 8.8 989 1 US-08-070-301-14 Sequence 14, Appli
38 70.5 8.8 989 1 US-09-949-016-9905 Sequence 9905, Ap
39 69.5 8.7 608 2 US-09-949-016-6888 Sequence 6888, Ap
40 69.5 8.7 823 2 US-08-476-515A-12 Sequence 12, Appli
41 69 8.6 943 2 US-08-476-515A-84 Sequence 84, Appli
42 69 8.6 4654 2 US-09-393-245-2 Sequence 5579, Ap
43 68.5 8.6 432 2 US-09-134-000C-5579 Sequence 5579, Ap
44 68.5 8.6 1206 2 US-08-469-260A-81 Sequence 81, Appli
45 67.5 8.4 1422 2

ALIGNMENTS

RESULT 1

US-09-621-976-3903
; Sequence 3903, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3903
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-621-976-3903

Query Match 100.0%; Score 799; DB 2; Length 153;
Best Local Similarity 100.0%; Pred No. 9,6e-94;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHLMGCPAN 60
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLYIFADELHLMGCPAN 60
QY 61 RIHTVYVEFIYLRVCGIRTVVSEETLLFQTELYETPRNIDHDPOEIHLESTSRKSW 120
Db 61 RIHTVYVEFIYLRVCGIRTVVSEETLLFQTELYETPRNIDHDPOEIHLESTSRKSW 120
QY 121 LTPVSTENEIKLDPSPFIADFTTAEELGLLIF 153
Db 121 LTPVSTENEIKLDPSPFIADFTTAEELGLLIF 153

RESULT 2

US-09-621-976-5222
; Sequence 5222, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2

```
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5222
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-621-976-5222
```

```
Query Match          98.7%; Score 789; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.9e-92;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYVEFIYLRDCGIRTRVVSSETLLFQTELYFTPRNIDHDPQEIHLECSRSKSVW 120
Db 61 RIHTVYVEFIYLRDCGIRTRVVSSETLLFQTELYFTPRNIDHDPQEIHLECSRSKSVW 120

QY 121 LTPVSTENEIKLDPSPFIADFQTAAELGLL 151
Db 121 LTPVSTENEIKLDPSPFIADFQTAAELGLL 151
```

```
RESULT 3
US-10-104-047-3674
; Sequence 3674, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674
```

```
Query Match          98.2%; Score 785; DB 2; Length 158;
Best Local Similarity 99.3%; Pred. No. 6.1e-92;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMASVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYVEFIYLRDCGIRTRVVSSETLLFQTELYFTPRNIDHDPQEIHLECSRSKSVW 120
Db 61 RIHTVYVEFIYLRDCGIRTRVVSSETLLFQTELYFTPRNIDHDPQEIHLECSRSKSVW 120

QY 121 LTPVSTENEIKLDPSPFIADFQTAAELGLL 151
Db 121 LTPVSTENEIKLDPSPFIADFQTAAELGLL 151
```

```
RESULT 4
US-09-621-976-5223
; Sequence 5223, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
```

```
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5223
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-621-976-5223
```

```
Query Match          53.7%; Score 429; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 9e-47;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYVEFIYLRDCGIRTRV 82
Db 61 RIHTVYVEFIYLRDCGIRTRV 82
```

```
RESULT 5
US-09-949-016-11232
; Sequence 11232, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11232
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11232
```

```
Query Match          28.0%; Score 223.5; DB 2; Length 216;
Best Local Similarity 33.3%; Pred. No. 4.3e-20;
Matches 40; Conservative 32; Mismatches 43; Indels 5; Gaps 2;
```

```
QY 5 VMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPANR1HT 64
Db 13 LMLLTSAFSAGSGQSPMTVLCSDIDWFMVTVHPFMLNNDVCVHPFELHGLGCPNHPVQ 72

QY 65 VYVEFIYLRDCGIRTRVVSSETLLFQTELYFTPRNIDHDPQEIHLECSRSKSVWLT 122
Db 73 HAYQFTYRVTGIRAKAVSQDMVIYSTEIHYSSKG---TPSKFVIVSCAAPQKSPWLT 129
```

```
RESULT 6
US-09-621-976-6447
; Sequence 6447, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
```

```

QY 79 --RTRVSEETLLFOTELYFTPRN--IDHDPQEIHLECSSTKRSV-----WLT-----PV 124
DB 132 PLRFLIIANTMAFQNDVYEWARHRAHKFSETHADPHNSRRGFFFSHVGCWLLVVRKHQA 191
QY 125 STENEIKLDSPFIAD-----FQTABELGLLI 152
DB 192 VKKGGKLDMSDLKAELVMFORRYKPGLLL 223

RESULT 8
US-09-723-806A-8
; Sequence 8, Application US/09723806A
; Patent No. 6686185
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Logan, Thomas Joseph
; TITLE OF INVENTION: 25934, A NOVEL FATTY ACID DESATURASE AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 10448-024001
; CURRENT APPLICATION NUMBER: US/09/723,806A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,455
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-723-806A-8

Query Match          9.6%; Score 77; DB 2; Length 358;
Best Local Similarity 25.0%; Pred. No. 0.44;
Matches 38; Conservative 22; Mismatches 50; Indels 42; Gaps 8;

QY 42 RNLYIFADELHLG-----MGPANRIHTYVYE-FIYLVRDCGI----- 78
DB 73 RNIILMA-LRHVGALYGITILPSSKVYLLMGIFYYLISALGITAGAHRLWSHRTYKARL 131
QY 79 --RTRVSEETLLFOTELYFTPRN--IDHDPQEIHLECSSTKRSV-----WLT-----PV 124
DB 132 PLRFLIIANTMAFQNDVYEWARHRAHKFSETHADPHNSRRGFFFSHVGCWLLVVRKHQA 191
QY 125 STENEIKLDSPFIAD-----FQTABELGLLI 152
DB 192 VKKGGKLDMSDLKAELVMFORRYKPGLLL 223

RESULT 9
US-09-583-110-3248
; Sequence 3248, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3248
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3248

```

Query Match 9.5%; Score 76; DB 2; Length 158;
Best Local Similarity 24.7%; Pred. No. 0.18;
Matches 39; Conservative 27; Mismatches 46; Indels 46; Gaps 9;

QY 22 VKISCSLWLMVSVIPVAE-----SRNLVIFADELHLMG--CPANRIHTYVYEF----- 69
DB 11 VVISTIDYLIILILFAQLSQNKQKWHIYAGQ-YLGTGLLVGASLVAAVYVNFVPEWM 69

QY 70 -----IYLVRDGGIRTRVVSSETLLFQTELYFTPRNIDHPOEI--HLECSSTRKSV 119
DB 70 VLLGLIPIYL-----GIRFAIVGSDA-----EEEEIEIERLEQSKANQLF 111

QY 120 W----LTPVSTENEIKLDPSPFIADFOQTABELGLLIF 153
DB 112 WTVTLTIASGDNLGI-YIPVFASLDWSQTLVALLVF 148

RESULT 10
US-09-107-433-3954
; Sequence 3954, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3954:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...165
; SEQUENCE DESCRIPTION: SEQ ID NO: 3954:
US-09-107-433-3954

Query Match 9.5%; Score 76; DB 2; Length 165;
Best Local Similarity 24.7%; Pred. No. 0.19;
Matches 39; Conservative 27; Mismatches 46; Indels 46; Gaps 9;

QY 22 VKISCSLWLMVSVIPVAE-----SRNLVIFADELHLMG--CPANRIHTYVYEF----- 69
DB 18 VVISTIDYLIILILFAQLSQNKQKWHIYAGQ-YLGTGLLVGASLVAAVYVNFVPEWM 76

QY 70 -----IYLVRDGGIRTRVVSSETLLFQTELYFTPRNIDHPOEI--HLECSSTRKSV 119
DB 77 VLLGLIPIYL-----GIRFAIVGSDA-----EEEEIEIERLEQSKANQLF 118

QY 120 W----LTPVSTENEIKLDPSPFIADFOQTABELGLLIF 153
DB 119 WTVTLTIASGDNLGI-YIEYFASLDWSQTLVALLVF 155

RESULT 11
US-09-489-039A-10162
; Sequence 10162, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10162
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10162

Query Match 9.5%; Score 76; DB 2; Length 323;
Best Local Similarity 24.3%; Pred. No. 0.51;
Matches 34; Conservative 23; Mismatches 55; Indels 28; Gaps 6;

QY 14 WTGAENLHVKISCSLWLMVSVIPVAES--RNLYIFADELHLMG-----CPANR-- 61
DB 149 WMSADNALLYVDNIRDALVKYDPPHADTVRNAAEYKEKIRQTMAPLQARLAQLPADKRW 208

QY 62 IHTYVYEFIYLVRDGGIRTRVVSSETLLFQTELYFTPRNIDHD--PQIHLECSSTRKSV 119
DB 209 LVTSEGAFSYLARDYGLR-----ELXLPINADQQGTPQVRKVIDTMRKE- 254

QY 120 WLPVSTENEIKLDPSPFIA 139
DB 255 RIPTIFSESTISDKPARQVA 274

RESULT 12
US-09-328-352-7887
; Sequence 7887, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7887
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7887

Query Match 9.4%; Score 75.5; DB 2; Length 385;
Best Local Similarity 25.0%; Pred. No. 0.76;
Matches 31; Conservative 16; Mismatches 52; Indels 25; Gaps 4;

QY 44 LVYFADELHLMGCPANRIHTYVYEFIYLVRDGGIRTRVVSSETL-----L 89

Db 75 LYLLELHI-----AQKDEFAIQOLSHIRSLGNEITAAQAEIRQKEYESSRQPDALDF 129
Qy 90 FOTELYTPRNIDHPQEIHECSTSRKSV-----WLPVSTENEIKLDPSPFTADF-QT 143
Db 130 POAQTYEPPKNTDTTAQFDELTTSSSEASFNDLKDYTPVKQEPAVEIEPLEFNFSEQT 189
Qy 144 TAE 147
Db 190 SAT 193

RESULT 13
US-08-484-993B-18
; Sequence 18, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-993B-18

Query Match 9.4%; Score 75.5; DB 1; Length 424;
Best Local Similarity 22.6%; Pred. No. 0.88;
Matches 42; Conservative 27; Mismatches 50; Indels 67; Gaps 11;
Qy 1 MALEVLMLLAVLIWTGA-----ENLH-----VKISCSLDLWMSVIPVAESR 42
Db 1 MGLSYGLFCFLWAGTGLCYPTTTEDKTHSLPSSPVVVECRHAWLVNV-----SK 55
Qy 43 NLY-----IFADELHGMGCPAN-----RIHTYVYEFYLVRCGIRPVVS 84
Db 56 NLFGTGLRVPADLTGLG---PENCEPLISGSDSDTTRFEVELHK-----CGNSVQ-VT 104
Qy 85 EBTLLFQTELYTPRNIDH-----DPQEIHECSTSRKSVMLTPVSTENEIKLDPSPFI 138

Db 105 EDALVSTFLHNPMPGNLSILTRAEVPIECRYPRHS-----NVSSEAIL-----PTW 155
Qy 139 ADFQTT 144
Db 156 VPFRTT 161
RESULT 14
US-08-484-158B-18
; Sequence 18, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-158B-18

Query Match 9.4%; Score 75.5; DB 1; Length 424;
Best Local Similarity 22.6%; Pred. No. 0.88;
Matches 42; Conservative 27; Mismatches 50; Indels 67; Gaps 11;
Qy 1 MALEVLMLLAVLIWTGA-----ENLH-----VKISCSLDLWMSVIPVAESR 42
Db 1 MGLSYGLFCFLWAGTGLCYPTTTEDKTHSLPSSPVVVECRHAWLVNV-----SK 55
Qy 43 NLY-----IFADELHGMGCPAN-----RIHTYVYEFYLVRCGIRPVVS 84
Db 56 NLFGTGLRVPADLTGLG---PENCEPLISGSDSDTTRFEVELHK-----CGNSVQ-VT 104

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:15:27 ; Search time 66.6 Seconds
(without alignments)
959.878 Million cell updates/sec

Title: US-10-664-025-3903
Perfect score: 799
Sequence: 1 MALEVLMLAVLIWTGAENL.....PSPFIADFTTAEELGLLIP 153

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	789	98.7	158	3	US-09-731-872-260
2	789	98.7	158	3	US-09-876-997-260
3	789	98.7	158	4	US-10-050-704-113
4	789	98.7	158	4	US-10-467-535-4
5	789	98.7	158	4	US-10-798-512-113
6	789	98.7	158	5	US-10-467-046-3
7	789	98.7	158	5	US-10-643-836-260
8	789	98.7	158	5	US-10-472-533-464
9	789	98.2	158	4	US-10-104-047-3674
10	429	53.7	106	3	US-09-731-872-265
11	429	53.7	106	3	US-09-876-997-265
12	429	53.7	106	5	US-10-643-836-265
13	225.5	28.2	212	4	US-10-287-971-40
14	223.5	28.0	212	3	US-09-745-763-203
15	223.5	28.0	212	4	US-10-047-021-91
16	223.5	28.0	212	5	US-10-970-493-91
17	223.5	28.0	213	4	US-10-411-224-91
18	161	20.2	163	4	US-10-287-971-42
19	98.5	12.3	930	5	US-10-450-763-42407
20	94	11.8	59	3	US-09-864-761-36906
21	88.5	11.1	234	5	US-10-450-763-57701
22	85.5	10.7	5373	4	US-10-408-765A-741
23	85.5	10.7	5430	5	US-10-805-684-151
24	84	10.5	1241	4	US-10-282-122A-46016
25	77	9.6	358	4	US-10-377-072-14
26	77	9.6	358	4	US-10-377-072-14
27	76	9.5	165	5	US-10-617-320-3954

28	76	9.5	204	5	US-10-472-928-3336	Sequence 3336, Ap
29	76	9.5	2957	5	US-10-732-923-8692	Sequence 8692, Ap
30	74.5	9.3	384	4	US-10-369-493-2111	Sequence 2111, Ap
31	74.5	9.3	525	4	US-10-425-115-253855	Sequence 253855,
32	73	9.1	376	4	US-10-087-192-9	Sequence 9, Appli
33	72	9.0	358	4	US-10-205-331-42	Sequence 42, Appl
34	71.5	8.9	318	4	US-10-306-762-201	Sequence 201, App
35	71.5	8.9	327	4	US-10-282-122A-58245	Sequence 58245, A
36	71.5	8.9	340	5	US-10-450-763-40925	Sequence 40925, A
37	71.5	8.9	649	4	US-10-104-047-3492	Sequence 3492, Ap
38	71	8.9	580	4	US-10-437-963-166460	Sequence 166460,
39	71	8.9	4796	6	US-11-097-143-2787	Sequence 2787, Ap
40	70.5	8.8	547	4	US-10-424-599-231071	Sequence 231071,
41	70	8.8	379	6	US-11-097-143-12645	Sequence 12645, A
42	70	8.8	1255	5	US-10-876-086-42	Sequence 42, Appl
43	69.5	8.7	291	6	US-11-097-143-30285	Sequence 30285, A
44	69.5	8.7	311	4	US-10-437-963-125990	Sequence 125990,
45	69.5	8.7	429	3	US-09-745-763-205	Sequence 205, App

ALIGNMENTS

RESULT 1
US-09-731-872-260
; Sequence 260, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 260
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-731-872-260

Query Match 98.7%; Score 789; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALEVLMLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLVIFADELHGMGCPAN 60
DB 1 MALEVLMLAVLIWTGAENLHVKISCSLDWLVSVIPVAESRNLVIFADELHGMGCPAN 60
QY 61 RHTTVYEFYLVRCGIRTVVSVETLLFQTELYFTPRNIDHDPOEIHLECSRSKSVW 120
DB 61 RHTTVYEFYLVRCGIRTVVSVETLLFQTELYFTPRNIDHDPOEIHLECSRSKSVW 120
QY 121 LTPVSTENEIKLDPSPFIADFTTAEELGLL 151
DB 121 LTPVSTENEIKLDPSPFIADFTTAEELGLL 151

RESULT 2
US-09-876-997-260
; Sequence 260, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste

APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US4.CIP
CURRENT APPLICATION NUMBER: US/09/876,997
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 09/731,872
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 260
LENGTH: 158
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -17...-1
US-09-876-997-260

Query Match 98.7%; Score 789; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60

Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYEFIVLRDCGIRTVVSEITLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120

Db 61 RIHTVYEFIVLRDCGIRTVVSEITLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 61 RIHTVYEFIVLRDCGIRTVVSEITLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
Db 61 RIHTVYEFIVLRDCGIRTVVSEITLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151
Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

RESULT 4

US-10-467-535-4
Sequence 4, Application US/10467535
Publication No. US20040146970A1
GENERAL INFORMATION:
APPLICANT: YUE, Henry; YAO, Monique G.
APPLICANT: ISON, Craig H.; LU, Yan
APPLICANT: WARREN, Bridget A.; ELLIOTT, Vicki S.
APPLICANT: BAUGHN, Mariah R.; DING, Li
APPLICANT: XU, Yuming; GIETZEN, Kimberly J.
APPLICANT: TANG, Tom Y.; LAL, Preeti G.
APPLICANT: DUGGAN, Brendan M.; BURFORD, Neil
APPLICANT: LU, Dyung Aina M.; RICHARDSON, Thomas W.
APPLICANT: TRAN, Uyen K.; KHARE, Reena
APPLICANT: CHAWLA, Narinder K.
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
FILE REFERENCE: PF-0903 USN
CURRENT APPLICATION NUMBER: US/10/467,535
CURRENT FILING DATE: 2003-08-08
PRIOR APPLICATION NUMBER: PCT/US02/03715
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/268,111
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/271,175
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: US 60/274,503
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US 60/274,552
PRIOR FILING DATE: 2001-03-09
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 158
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 5284076CD1
US-10-467-535-4

Query Match 98.7%; Score 789; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60

Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60

QY 61 RIHTVYEFIVLRDCGIRTVVSEITLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120

Db 61 RIHTVYEFIVLRDCGIRTVVSEITLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151

RESULT 5

US-10-798-512-113
Sequence 113, Application US/10798512
Publication No. US20040152164A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.


```
; SEQ ID NO 464
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-472-533-464

Query Match          98.7%; Score 789; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 RIHTVYVEFIYLRDCGIRTVVSEBTLFQTELYFTPRNIDHDPOEIHLECSRKSVW 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 RIHTVYVEFIYLRDCGIRTVVSEBTLFQTELYFTPRNIDHDPOEIHLECSRKSVW 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 LTPVSTENEIKLDPSPFIADFTTAEELGLL 151
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 LTPVSTENEIKLDPSPFIADFTTAEELGLL 151
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9
US-10-104-047-3674
; Sequence 3674, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3674
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3674

Query Match          98.2%; Score 785; DB 4; Length 158;
Best Local Similarity 99.3%; Pred. No. 8.1e-84;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMASVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 RIHTVYVEFIYLRDCGIRTVVSEBTLFQTELYFTPRNIDHDPOEIHLECSRKSVW 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 RIHTVYVEFIYLRDCGIRTVVSEBTLFQTELYFTPRNIDHDPOEIHLECSRKSVW 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 LTPVSTENEIKLDPSPFIADFTTAEELGLL 151
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 LTPVSTENEIKLDPSPFIADFTTAEELGLL 151
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 10
US-09-731-872-265
; Sequence 265, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
```

```
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-731-872-265

Query Match          53.7%; Score 429; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.6e-42;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 RIHTVYVEFIYLRDCGIRTV 82
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 RIHTVYVEFIYLRDCGIRTV 82
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 11
US-09-876-997-265
; Sequence 265, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-09-876-997-265

Query Match          53.7%; Score 429; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.6e-42;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MALEVLMLLAVLIWTGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHGMGCPAN 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 RIHTVYVEFIYLRDCGIRTV 82
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 RIHTVYVEFIYLRDCGIRTV 82
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 12
US-10-643-836-265
; Sequence 265, Application US/10643836
; Publication No. US20050096458A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/10/643,836
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 265
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-10-643-836-265

```

Query Match	53.7%	Score 429;	DB 5;	Length 106;
Best Local Similarity	100.0%;	Pred.No. 3.6e-42;		
Matches	82;	Conservative	0; Mismatches	0; Indels
				0; Gaps
Qy	1	MALEVLMLLAVLTWTGAENLHVKISCSLDWLMVSVI	PAESRNLYIFADELHGMGCPAN	60
Dd	1	MALEVLMLLAVLTWTGAENLHVKISCSLDWLMVSVI	PAESRNLYIFADELHGMGCPAN	60
Qy	61	RIHTYVVEFYILVRDCGIRTV	82	
Dd	61	RIHTYVVEFYILVRDCGIRTV	82	

```

RESULT 13
US-10-287-971-40
; Sequence 40, Application US/10287971
; Publication No. US20040067882A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook, et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-480A
; CURRENT APPLICATION NUMBER: US/10/287,971
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 09/997,425
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 10/035,568
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/338,626
; PRIOR FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/401,479
; PRIOR FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: 60/333,072
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/348,283
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/393,262
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/406,181
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 397
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 40
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-971-40

```

Query Match
28.2%; Score 225.5; DB 4; Length 212;

	Best Local Similarity Matches	33.3%; Conservative	Pred. No. 7.9e-18; 33; Mismatch	Indels	Gaps
QY	5	VLMLLAVLIWTGAENLHKISCSLDLWLSVIPVAESNLIYFADELHGLMGCPNRRVHT	64		
DB	9	LMILLTALSAGSQSPMTVLCISIDWFNVTHPMLNNDVCVPHLHGLGCPNNHVQP	68		
QY	65	YVYEFIVLRDCGTRIRVVSEETLLFOTELVFTPRNIDHDPOE--IHLECSSTRSKSVMLT	122		
DB	69	HAYQFVTRVTECGIRAKAVSDMWIYSTEIHYSKSG---TPSKFVIVPSCAAPQSKSPMLT	125		

RESULT 14

```

US-09-745-763-203
; Sequence 203, Application US/09745763
; Patent No. US20020065394A1
GENERAL INFORMATION:
;
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; LaVallie, Edward R.
; Collins-Racie, Lisa A.
; Evans, Cheryl
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
;
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; ENCODING THEM

```

NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/745,763

```

;
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

```

, NAME: Springer, Suzanne A.
, REGISTRATION NUMBER: 41,323
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (617) 498-8284
, TELEFAX: (617) 876-5851
, INFORMATION FOR SEQ ID NO: 203:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 212 amino acids
, TYPE: amino acid
, STRANDEDNESS: <Unknown>
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, SEQUENCE DESCRIPTION: SEQ ID NO:
US-09-745-763-203

```

		Query Match	28.0%;	Score 223.5;	DB 3;	Length 212;		
		Best Local Similarity	33.3%;	Pred. No. 1.4e-17;				
Matches	40;	Conservative	32;	Mismatches	43;	Indels	5; Gaps	2;
<hr/>								
Qy	5	VLMMLAVLIWTGAENLHVKISCSLDWIMVSVIPVAESRNLYIFADELHLGMGCPANRIHT	64					
	:	:::::	: :	:	:	:	:	:
Dd	9	LMIILTSAFSAGSQSPMTVLCSIDWMFTVTVPFMNLNDVCVFHFLHELGLGCPPNHQP	68					
	:	:	:	:	:	:	:	:
<hr/>								
Qy	65	YVEFYIYLVDCCIRTVRVSEETLLFQTLEYLPFRNIDHPQB--IHLECSTSRKSVALT	122					
	:	:	:	:	:	:	:	:
Dd	69	HAYQFYRYRTECGIRAKAVQSODMVIYSTEIHYSGK---TPSKFVIPVSCAAPOKSPMLT	125					
	:	:	:	:	:	:	:	:

Wed May 10 09:24:02 2006

```
RESULT 15
US-10-047-021-91
; Sequence 91, Application US/10047021
; Publication No. US20040002591A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: PZ016P2
; CURRENT APPLICATION NUMBER: US/10/047,021
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,066
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: PCT/US98/18360
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/090,112
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 91
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any amino acid
US-10-047-021-91

Query Match      28.0%; Score 223.5; DB 4; Length 212;
Best Local Similarity 33.3%; Pred. No. 1.4e-17;
Matches 40; Conservative 32; Mismatches 43; Indels 5; Gaps 2;

QY      5  VLMLLAVLIWTGAENLHVKISCLDMLMVSVIPVAESRNLYIFADELHGMGCPANRIHT 64
Db      9  LMLITSAPSGSGSPMTVLCSIDFWFVTHVHFFMLNNDVCVHFHFLHGLGCPNNHVQP 68

QY      65  VYVEFIYLRDCCGIRTRVVEETLTFOTELYTPRNIDHPQE--IHLECSSTRKSVMLT 122
Db      69  HAYQFTYRTECGIRAKAVSQDMVIYSTEIHVSSKG---TPSKFVIPVSCAAPQKSPWLT 125

Search completed: May 9, 2006, 10:17:00
Job time : 67.6 secs
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006. Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:15:47 ; Search time 19.8 Seconds
(without alignments)
357.655 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLWITGAENL.....PSPFIADFQTAEELGLLIF 153

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

235405 seqs, 46284737 residues

Total number of hits satisfying chosen parameters: 235405

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:
1: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep:1*
2: /SID55/ptodata/1/pubpaa/US06_NEW_PUB.pep:2*
3: /SID55/ptodata/1/pubpaa/US07_NEW_PUB.pep:3*
4: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep:4*
5: /SID55/ptodata/1/pubpaa/PCT_NEW_PUB.pep:5*
6: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep:6*
7: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep:7*
8: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep:8*
9: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep:9*
10: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep:10*
11: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep:11*
12: /SID55/ptodata/1/pubpaa/US60_NEW_PUB.pep:12*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	785	98.2	158	11	US-11-072-512-3674
2	775	97.0	148	9	US-10-475-075-788
3	752	94.1	143	9	US-10-475-075-787
4	223.5	28.0	216	9	US-10-821-234-1033
5	85.5	10.7	3803	9	US-10-995-561-773
6	85.5	10.7	3960	9	US-10-995-561-771
7	85.5	10.7	5335	9	US-10-995-561-777
8	85.5	10.7	5406	9	US-10-995-561-774
9	85.5	10.7	5415	9	US-10-995-561-779
10	85.5	10.7	5464	9	US-10-995-561-775
11	76.5	9.6	242	11	US-11-096-568A-8621
12	76.5	9.6	394	11	US-11-096-568A-8620
13	76.5	9.6	398	11	US-11-096-568A-8619
14	73	9.1	460	11	US-11-188-298-7840
15	71.5	8.9	649	11	US-11-072-512-3492
16	67.5	8.4	393	11	US-11-188-298-11630
17	67.5	8.4	477	11	US-11-098-686-10915
18	67	8.4	269	11	US-11-079-463-8166
19	66.5	8.3	440	11	US-11-045-004-81
20	66	8.3	497	11	US-11-188-298-15337
21	65.5	8.2	489	9	US-10-835-475-11

22	65.5	8.2	514	9	US-10-835-475-2	Sequence 2, Appli
23	65.5	8.2	1032	9	US-10-835-475-1	Sequence 1, Appli
24	65.5	8.2	1032	9	US-10-194-487-552	Sequence 552, App
25	65.5	8.2	1032	9	US-10-195-883-552	Sequence 552, App
26	65.5	8.2	1032	9	US-10-195-888-552	Sequence 552, App
27	65.5	8.2	1032	9	US-10-195-889-552	Sequence 552, App
28	65.5	8.2	1032	11	US-11-014-367-1	Sequence 1, Appli
29	65.5	8.2	1032	11	US-11-185-567-2	Sequence 2, Appli
30	65	8.1	150	11	US-11-096-568A-28984	Sequence 28984, A
31	65	8.1	174	11	US-11-096-568A-28983	Sequence 28983, A
32	65	8.1	252	11	US-11-096-568A-28982	Sequence 28982, A
33	65	8.1	476	11	US-11-288-493-8	Sequence 8, Appli
34	64	8.0	315	11	US-11-098-686-11063	Sequence 11063, A
35	64	8.0	497	11	US-11-096-568A-5581	Sequence 5581, Ap
36	64	8.0	524	9	US-10-467-657-8122	Sequence 8122, Ap
37	64	8.0	524	11	US-11-188-298-3472	Sequence 3472, Ap
38	64	8.0	604	11	US-11-096-568A-5580	Sequence 5580, Ap
39	64	8.0	624	11	US-11-096-568A-5579	Sequence 5579, Ap
40	64	8.0	838	11	US-11-031-737A-11	Sequence 11, Appl
41	64	8.0	838	11	US-11-031-482-11	Sequence 11, Appl
42	64	8.0	1111	11	US-11-087-099-998	Sequence 998, App
43	64	8.0	1111	11	US-11-087-099-5147	Sequence 5147, Ap
44	63.5	7.9	359	11	US-11-188-298-5925	Sequence 5925, Ap
45	63.5	7.9	396	11	US-11-096-568A-34086	Sequence 34086, A

ALIGNMENTS

RESULT 1

US-11-072-512-3674

; Sequence 3674, Application US/11072512

; Publication No. US20060029945A1

; GENERAL INFORMATION:

; APPLICANT: ISOQAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAHIKO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOKYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: Novel full length cDNA

; FILE REFERENCE: 084335-0191

; CURRENT APPLICATION NUMBER: US/11/072,512

; CURRENT FILING DATE: 2005-03-07

; PRIOR APPLICATION NUMBER: US 60/350,978

; PRIOR FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: JP 2001-379298

; PRIOR FILING DATE: 2001-11-05

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 3674

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-072-512-3674

Query Match 98.2%; Score 785; DB 11; Length 158;

Best Local Similarity 99.3%; Pred. No. 3 6e-82;

Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLWITGAENLHVYKISCSLDLWLMVSVIPVAESRNLVYFADELHGMGCPAN 60

```
Db      1 MALEVLMLLAVLIWITGAENLHVKISCSLDWLMASVIPAESRNLYIFADELHLMGCPAN 60
QY      61 RIHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
Db      61 RIHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
QY      121 LTPVSTENEIKLDPSPFIADFQTTAAELGLL 151
Db      121 LTPVSTENEIKLDPSPFIADFQTTAAELGLL 151

RESULT 2
US-10-475-075-788
; Sequence 788, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
; SOFTWARE: Patent.pm
; SEQ ID NO 788
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-475-075-788

Query Match      97.0%; Score 775; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 4.6e-81;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWITGAENLHVKISCSLDWLMVSVIPAESRNLYIFADELHLMGCPAN 60
Db      1 MALEVLMLLAVLIWITGAENLHVKISCSLDWLMVSVIPAESRNLYIFADELHLMGCPAN 60
QY      61 RIHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
Db      61 RIHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
QY      121 LTPVSTENEIKLDPSPFIADFQTTAAEL 148
Db      121 LTPVSTENEIKLDPSPFIADFQTTAAEL 148

RESULT 3
US-10-475-075-787
; Sequence 787, Application US/10475075
; Publication No. US20060053498A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: Full-length human cDNAs encoding potentially secreted proteins
; FILE REFERENCE: G-081US03PCT
; CURRENT APPLICATION NUMBER: US/10/475,075
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/IB01/00914
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 918
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 787
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17...-1
US-10-475-075-787

Query Match      94.1%; Score 752; DB 9; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.9e-78;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALEVLMLLAVLIWITGAENLHVKISCSLDWLMVSVIPAESRNLYIFADELHLMGCPAN 60
Db      1 MALEVLMLLAVLIWITGAENLHVKISCSLDWLMVSVIPAESRNLYIFADELHLMGCPAN 60
QY      61 RIHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
Db      61 RIHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOEIHLECSTSRKSVW 120
QY      121 LTPVSTENEIKLDPSPFIADFQT 143
Db      121 LTPVSTENEIKLDPSPFIADFQT 143

RESULT 4
US-10-821-234-1033
; Sequence 1033, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1033
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1033

Query Match      28.0%; Score 223.5; DB 9; Length 216;
Best Local Similarity 33.3%; Pred. No. 6.8e-18;
Matches 40; Conservative 32; Mismatches 43; Indels 5; Gaps 2;

QY      5 VLMLLAVLIWITGAENLHVKISCSLDWLMVSVIPAESRNLYIFADELHLMGCPANRIHT 64
Db      13 LMLLITSAFSAGSGQSPMTVLCSIDFWVTWVHPFPMNLNDVCVHPFHELHGLGCPNHPQP 72
QY      65 VYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPOE--IHLECSTSRKSVMLT 122
Db      73 HAYQTYRVTGCGIRAKAVSQDMVIYSTEIHSYSGK--TPSKFVIPVSCAAPQKSPMLT 129

RESULT 5
US-10-995-561-773
; Sequence 773, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
```

```
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 773
; LENGTH: 3803
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-773

Query Match      10.7%; Score 85.5; DB 9; Length 3803;
Best Local Similarity 25.4%; Pred. No. 1.5;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCSLDMLMVSVIPVAESRNLYIFADELHLMGCPA--- 59
Db 426 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQPVOCE 470

QY 60 NRHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCEST-SRK- 117
Db 471 SDVIMYIOCEGLIRQLQVDLQILRDENYYQLEELAFVMRLQDELVTLRLECTNLYRKG 530

QY 118 ---SVMLTPVS--TENEIKLDP 134
Db 531 HFTSLELVPPSTLTTHLKAEP 552

RESULT 6
US-10-995-561-771
; Sequence 771, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 771
; LENGTH: 3960
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-771

Query Match      10.7%; Score 85.5; DB 9; Length 3960;
Best Local Similarity 25.4%; Pred. No. 1.6;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCSLDMLMVSVIPVAESRNLYIFADELHLMGCPA--- 59
Db 426 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQPVOCE 470

QY 60 NRHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCEST-SRK- 117
Db 471 SDVIMYIOCEGLIRQLQVDLQILRDENYYQLEELAFVMRLQDELVTLRLECTNLYRKG 530

QY 118 ---SVMLTPVS--TENEIKLDP 134
Db 531 HFTSLELVPPSTLTTHLKAEP 552

RESULT 7
US-10-995-561-777
; Sequence 777, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 777
; LENGTH: 5335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-777

Query Match      10.7%; Score 85.5; DB 9; Length 5335;
Best Local Similarity 25.4%; Pred. No. 2.3;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCSLDMLMVSVIPVAESRNLYIFADELHLMGCPA--- 59
Db 334 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQPVOCE 378

QY 60 NRHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCEST-SRK- 117
Db 379 SDVIMYIOCEGLIRQLQVDLQILRDENYYQLEELAFVMRLQDELVTLRLECTNLYRKG 438

QY 118 ---SVMLTPVS--TENEIKLDP 134
Db 439 HFTSLELVPPSTLTTHLKAEP 460

RESULT 8
US-10-995-561-774
; Sequence 774, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 774
; LENGTH: 5406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-774

Query Match      10.7%; Score 85.5; DB 9; Length 5406;
Best Local Similarity 25.4%; Pred. No. 2.3;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCSLDMLMVSVIPVAESRNLYIFADELHLMGCPA--- 59
Db 426 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQPVOCE 470

QY 60 NRHTVYVEFIYLRDCGIRTVVSEETLLFQTELYFTPRNIDHDPQEIHLCEST-SRK- 117
Db 471 SDVIMYIOCEGLIRQLQVDLQILRDENYYQLEELAFVMRLQDELVTLRLECTNLYRKG 530

QY 118 ---SVMLTPVS--TENEIKLDP 134
Db 531 HFTSLELVPPSTLTTHLKAEP 552

RESULT 9
US-10-995-561-779
; Sequence 779, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
```

```

; ; TITLE OF INVENTION: DETECTION AND USES THEREOF
; ; FILE REFERENCE: CL001559
; ; CURRENT APPLICATION NUMBER: US/10/995,561
; ; CURRENT FILING DATE: 2004-11-24
; ; NUMBER OF SEQ ID NOS: 85702
; ; SOFTWARE: FastSEQ for Windows Version 4.0
; ; SEQ ID NO 779
; ; LENGTH: 5415
; ; TYPE: PRT
; ; ORGANISM: Homo sapiens
; ; US-10-995,561-779

```

Query Match	10.7%;	Score 85.5;	DB 9;	Length 5415;
Best Local Similarity	25.4%;	Pred. No. 2.3;		
Matches	36;	Conservative 23;	Mismatches 58;	Indels 25; Gaps 5;
Qy	3	LEVLMLAVLITGTAENLHVKSISCSLDLWLMVSVTPVAESRNLYIFADELHLMGCPA---	59	
Db	426	LELLQANKIONGALCEEKLTUAKNTLQ-----ADAAHLESQPVQCE	470	
Qy	60	NRIHTYVYFPIYLVRDCGIRTVVSEETLLFQTELYFTPRNIDHPDQFIHLECSY-SRK-	117	
Db	471	SDVIMYIQECEGLRIQVLQILRDENYQYLEELAFRVWELQDELVTYLRLECTNLYRKG	530	
Qy	118	---SVWLTPVS--TENEIKLDP	134	
Db	531	HFTSLELVPPSTLTTHLKAEP	552	

```

RESULT 10
US-10-995-561-775
; Sequence 775, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS IN THE CARDIOVASCULAR SYSTEM
; TITLE OF INVENTION: CARDIOVASCULAR SYSTEM
; TITLE OF INVENTION: DETECTION AND USE OF CARDIOVASCULAR SYSTEM
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 1.0
; SEQ ID NO 775
; LENGTH: 5464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-775

```

Query Match	10.7%	Score 85.5;	DB 9;	Length 5464;
Best Local Similarity	25.4%	Pred. No.2,3;		
Matches	36;	Conservative 23;	Mismatches 58;	Indels 25; Gaps 5;
Qy	3	LEVLMLAVLILWTCAENLHWKISCSLDLWLVSVIPVAESRNLYIFADELHGLMGCPA---	59	
Db	426	LELLQANKIQNGALNCEEKLTAKNTLQ-----ADAHLSESGPVOCE	470	
Qy	60	NRIHTYVVEFYLVDRDGCIRTRVVSSEETLLFQTELYFTPRNIDHPQSIHLECSST-SRK-	117	
Db	471	SDVIMYIQECEGLRIQVLDQILRDENYYQLESLAFRVVWLQDELVTYLRLECTNLYRKG	530	
Qy	118	---SVWLTPVS--TENEIKLDP	134	
Db	531	HFTSLELVPPSTLTTHLKAEP	552	

```

RESULT 11
US-11-096-568A-8621
; Sequence 8621, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined

```

```

; TITLE OF INVENTION:  Therby
; FILE REFERENCE:  2750-1592PUS2
; CURRENT APPLICATION NUMBER:  US/11/096,568A
; CURRENT FILING DATE:  2005-04-01
; NUMBER OF SEQ ID NOS:  34471
; SEQ ID NO 8621
; LENGTH:  242
; TYPE:  PRT
; ORGANISM:  Glycine max
; FEATURES:
; NAME/KEY:  misc_feature
; LOCATION:  (1)..(242)
; OTHER INFORMATION:  Ceres Seq. ID no. 15169117
US-11-096-568A-8621

```

	Query Match	9.6%	Score 76.5;	DB 11;	Length 242;
	Best Local Similarity	21.5%;	Pred. No. 0.48;		
	Matches 40;	Conservative 29;	Mismatches 70;	Indels 47;	Gaps 9
Qy	9	LAVLIWTGAENLH----	VKISCSLDWLMV-----	SVIPVAESRNLXYF----	ADELHLGM 55
Dd	21	LTALWFSDGSHDEMKILSLLSQHLLHENVADSYLPASSNMWYFSKDQASNYQGF 80	:	:	:
Qy	56	GCPSANRIHTVVYEFIVLRDCGRT-----	--RVSEETLLFOLEIYFTPRNIDHPD 105	:	:
Dd	81	NISWARLHVLYGD-IVYSQDSCESDVFELLVPRLVRDNELDDQFV-----	NIMHAL 134	:	:
Qy	106	QETHLEC-----	-SFRKSVMTLTPYST-----	NEIKLDPSFFIADF--QTAAEE 147	:
Dd	135	QVWHFYCWKGSAALLNNMSNTDSVLKVTEVQREKISGLNSVLLNDNFASKVCFCMFPSEE 194	:	:	:
Qy	148	LGLLIIF 153	:	:	:
Dd	195	KGCICF 200	:	:	:

```

RESULT 12
US-11-096-568A-8620
; Sequence 8620, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fr
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8620
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(394)
; OTHER INFORMATION: Ceres Seq. ID no. 15169116
US-11-096-568A-8620

```

	Query Match	9.6%;	Score 76.5;	DB 11;	Length 394;
	Best Local Similarity	21.5%;	Pred. No. 0.89;		
	Matches	40;	Conservative	29;	Mismatches 70; Indels 47; Gaps 9
Qy	9	LAVLIWTGAENLH---	VKISCSLDLMV----	SVIPVAESRNLYP----	ADELHQLQM 55
Db	173	LTALMFSRDSGHDEDKGLISLSQHLHENVADSYLPLPASSNNYWFQKQASNYQGF			232
Qy	56	GCPANRHTVYVEFIYLVRCGIRT-----			
Db	233	NISWARLHVLYGD-IYISQDSCESDYFELLVKPVLVRVDNELLDQFV-----			NIMHAL 286
Qy	106	QEIHLEC-----			
Db	287	OVVHFYCWKGSALNNMNSITDSRLVKTVEOREKISGLNSVLLANDNDPASKVFCNMPSEE			346

QY 148 LGLLIF 153
Db 347 KGCICF 352

RESULT 13

US-11-096-568A-8619
; Sequence 8619, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8619
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(398)
; OTHER INFORMATION: Ceres Seq. ID no. 15169115
US-11-096-568A-8619

Query Match 9.6%; Score 76.5; DB 11; Length 398;
Best Local Similarity 21.5%; Pred. No. 0.9;
Matches 40; Conservative 29; Mismatches 70; Indels 47; Gaps 9;

QY 9 LAVLIWTGAENLH---VKISCSLDLWLV-----SVIPVAESRNLIYF-----ADELHLMG 55
Db 177 LTLALMFSRGSDEDMKGLSLLSQHLHENVADSVLPLPASNMYWFSKDAQSNYQGF 236
QY 56 GCPANRIHTYVYEFILVRDCGIRT-----RWVSEETLLFOTELYFTPRNIDHDP 105
Db 237 NISWARLHVLYGD-IVISQSSCESDVPELLVPRLVRNELLDQGFV-----NIMHAL 290
QY 106 QETHLEC-----STSRKSVMLTPVSTE-----NEIKLDPSPFTADF--QTAAEE 147
Db 291 QVVHFYCWKGSALLNNMSNIDSVLKVTEVQREKISGLNSVLLNDNDFASKVFCNNPSEE 350
QY 148 LGLLIF 153
Db 351 KGCICF 356

RESULT 14

US-11-188-298-7840
; Sequence 7840, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 7840
; LENGTH: 460
; TYPE: PRT
; ORGANISM: GIBBERELLA ZEAE PH-1
US-11-188-298-7840

Query Match 9.1%; Score 73; DB 11; Length 460;
Best Local Similarity 22.8%; Pred. No. 2.7;
Matches 31; Conservative 21; Mismatches 36; Indels 48; Gaps 6;

QY 34 SVIPVAESRNLIYFADEL--HLGMGCPANRIHTYVYEFILVRDCGIRTVVSEETLLFQ 91

Db 115 SLVTAHSKNIIYVADVVANHMGKIQDHRPEP-----LNQ 150
QY 92 TELYFTPRNIDHDPQEIHLLEC-----STSRKSV-----WLTVPVSTE---NEIKLD- 133
Db 151 QSSVHSPCAIDYNNQNSIEQCEIADLPDLNTGSETVKVKLNDWISLWLVTEVSPDGIRDT 210
QY 134 -----PSPFIAFDQTTA 145
Db 211 VKHVEKSFWPDFQKAA 226

RESULT 15

US-11-072-512-3492
; Sequence 3492, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3492
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3492

Query Match 8.9%; Score 71.5; DB 11; Length 649;
Best Local Similarity 23.3%; Pred. No. 6.2;
Matches 38; Conservative 27; Mismatches 63; Indels 35; Gaps 7;

QY 9 LAVLIWTGAENLHVKISCSLDLWLVS-----VIPVAE-----SRNLIYFADE 50
Db 37 ISSICWSSNNFLVLTASSSGDKIVVSSCKKPVLLLEAEGQKQTCVNLNSTSMYLVSGG 96
QY 51 LHLGMG---CPANRIHTYVVE-----FIYLRDCGIRTVVSEETLLFQ---TELYFTP 98
Db 97 LANTVNIWDLKSKRVHRSLKDHKQDVTCVTYNNNDYIASGSLSGIILHSHVTNLSSTP 156
QY 99 RNIDHDPQBIHLECSTSRKSVMLTPVSTENEIKL-----DPSPF 137
Db 157 FGHGNSQSVRHLKCSLFFKSL-LGSVSDNGIVTLWDVNSQSPY 198

Search completed: May 9, 2006, 10:17:28
Job time : 19.8 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:13:22 ; Search time 17.1 Seconds
(without alignments)
860.887 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLIWTGAENL.....PSPFIADFTTAEELGLLIF 153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92.5	11.6	422	1 A60503	sperm-binding glycoprotein
2	84.5	10.6	1825	2 T42725	actin binding protein
3	84.5	10.6	1885	2 T30847	actin binding protein
4	84.5	10.6	2033	2 T30849	actin binding protein
5	78.5	9.8	135	2 G70519	hypothetical protein
6	77	9.6	358	1 A24699	stearoyl-CoA 9-des
7	76.5	9.6	367	2 A82094	thiamin biosynthesis
8	76.5	9.6	468	2 AE0152	probable bacteriophage
9	76	9.5	204	2 H95188	cadmium resistance
10	76	9.5	204	2 A98055	hypothetical protein
11	76	9.5	256	2 T45895	hypothetical protein
12	76	9.5	2957	2 T33152	hypothetical protein
13	75.5	9.4	424	2 S70399	zona pellucida glycoprotein
14	75	9.4	647	1 I58353	LIM protein kinase
15	74.5	9.3	232	2 H84028	DNA repair protein
16	74.5	9.3	424	1 A30334	sperm-binding glycoprotein
17	74.5	9.3	993	2 T38598	probable family 31
18	74	9.3	307	2 T23302	hypothetical protein
19	73	9.1	3036	2 T18995	hypothetical protein
20	72.5	9.1	1256	2 C71436	probable resistance
21	72	9.0	285	2 A84953	pantoate-beta-alan
22	71.5	8.9	327	2 H64155	conserved hypothetical
23	71.5	8.9	338	2 B64384	hypothetical protein
24	71.5	8.9	1188	2 S50434	hypothetical protein
25	71	8.9	424	1 A36000	sperm-binding glycoprotein
26	70.5	8.8	972	1 URBOAP	peptidylglycine monomeric
27	70.5	8.8	976	1 URBTAP	peptidylglycine monomeric
28	70	8.8	325	2 T25331	hypothetical protein
29	70	8.8	647	1 I48737	LIM protein kinase

30	69.5	8.7	513	2 P90488	long-chain-fatty-acid
31	69	8.6	2098	2 T25888	hypothetical protein
32	69	8.6	2467	2 D71437	probable resistance
33	68.5	8.6	458	2 E91092	probable invasion
34	68.5	8.6	458	2 A85938	hypothetical protein
35	68.5	8.6	464	2 T21628	hypothetical protein
36	68.5	8.6	514	2 AG2297	cobyrinic acid a,c
37	68	8.5	262	2 S78172	probable transport
38	68	8.5	273	2 D71436	hypothetical protein
39	68	8.5	336	2 AB2763	conserved hypothetical
40	68	8.5	336	2 A97544	nifH/smm1 family
41	68	8.5	355	2 A32115	stearoyl-CoA 9-des
42	68	8.5	499	2 E69123	glutamate synthase
43	68	8.5	503	1 DCFPD2	aromatic-L-amino-acid
44	68	8.5	510	1 DCFPA	aromatic-L-amino-acid
45	68	8.5	511	1 DCFPD1	aromatic-L-amino-acid

ALIGNMENTS

RESULT 1

A60503

sperm-binding glycoprotein ZP3 precursor - golden hamster

N;Alternate names: sperm receptor; zona pellucida glycoprotein ZP3

C;Species: Mesocricetus auratus (golden hamster)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: A60503

R;Kinloch, R.A.; Ruiz-Seiler, B.; Wassarman, P.M.

Dev. Biol. 142, 414-421, 1990

A;Title: Genomic organization and polypeptide primary structure of zona pellucida glycoprotein ZP3

A;Reference number: A60503; PMID:91078540; PMID:2257975

A;Accession: A60503

A;Molecule type: DNA

A;Residues: 1-422 <XIN>

A;Cross-references: UNIPARC:UPI00001744E3; GB:M63629

A;Note: the authors translated the codon CAA for residue 251 as Glu, and AGG for residue

C;Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a receptor

C;Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

C;Keywords: glycoprotein; oocyte

F;45-300/Domain: ZP domain homology <ZPH>

Query Match	11.6%	Score	92.5	DB	1	Length	422
Best Local Similarity	23.4%	Pred. No.	0.089				
Matches	51	Conservative	30	Mismatches	66	Indels	71
Gaps	12						
QY	1	MALEVLMLLAVLIWTGAENLHVKISGLD-WLMVSVIP-----VAE	40				
Db	1	MGLSVQLLCLLCCGA-----KQCCSQPLWLLPGGTPPGKLTSSVEVECLEAEIVVTV	55				
QY	41	SRNLY-----IFADELHGM-GC-PANRIHTYVYFIYLVRCGIRTVVSEETLLFQTE	93				
Db	56	SRDLFGTGLIQLPEDLTLSGSENCPLVSVATDVVRFKAQLHECSNRVQ-VTEALVYSTV	114				
QY	94	LYFTPRND-----HDPQEIHECSTSR-----KSVML---TPVSTENEI-----	130				
Db	115	LLHQPRPVGSLILNTRADVPICRYPQGNVSHAIRPTWVPFTTSSSEKLVFSLR	174				
QY	131	-----KLDPSPFIADFDQ-----TTAEELGLLIF	153				
Db	175	LMEENWNTEKLSPTSHLGEVAYLQAEVQVGTSHLPLLLF	212				

RESULT 2

T42725

actin binding protein ACF7, neural isoform 1 - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 03-Nov-2000

C;Accession: T42725

R;Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.

Genomics 38, 19-29, 1996

A;Title: Cloning and characterization of mouse ACF7, a novel member of the dystonin sub

A;Reference number: Z20900; PMID:97124842; PMID:8954775

A:Accession: T42725
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1825 <BER>
A:Cross-references: UNIPARC:UPI000016D118; EMBL:U67203; NID:g1675221; PID:g1675222; PIDN
C:Genetics:
A:Gene: ACF7
A:Map position: 4
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein 8
C:Keywords: actin binding

Query Match 10.6%; Score 84.5; DB 2; Length 1825;
Best Local Similarity 25.4%; Pred. No. 3.6;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCLDWLWVSVIPVAESRNLYIFADELHLGMGCPA--- 59
DB 369 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQFVQCE 413

QY 60 NRIHTYVVEFYLVNRDCGIRTVWSEETLLFOTELYFTPRNIDHDPQEIHLCEST-SRK- 117
DB 414 SDVIMYIQCEGLIRQLQVDQLRDEKYYQLEELAFVRMLQDELVTLRLECTNLYRKG 473

QY 118 ---SVMLTPVS--TENEIKLDP 134
DB 474 HFSSLELVPPSTLTTHLKAEP 495

RESULT 3
T30847
actin binding protein ACF7, neural isoform 2 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 03-Nov-2000
C:Accession: T30847
R:Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
Genomics 38, 19-29, 1996
A:Title: Cloning and characterization of mouse ACF7, a novel member of the dystonin subf
A:Reference number: Z20900; MUID:97124842; PMID:8954775
A:Accession: T30847
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1885 <BER>
A:Cross-references: UNIPARC:UPI000016D119; EMBL:U67204; NID:g1675223; PID:g1675224; PIDN
C:Genetics:
A:Gene: mACF7
A:Map position: 4
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein 8

Query Match 10.6%; Score 84.5; DB 2; Length 1885;
Best Local Similarity 25.4%; Pred. No. 3.8;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCLDWLWVSVIPVAESRNLYIFADELHLGMGCPA--- 59
DB 429 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQFVQCE 473

QY 60 NRIHTYVVEFYLVNRDCGIRTVWSEETLLFOTELYFTPRNIDHDPQEIHLCEST-SRK- 117
DB 474 SDVIMYIQCEGLIRQLQVDQLRDEKYYQLEELAFVRMLQDELVTLRLECTNLYRKG 533

QY 118 ---SVMLTPVS--TENEIKLDP 134
DB 534 HFSSLELVPPSTLTTHLKAEP 555

RESULT 4
T30849
actin binding protein ACF7, neural isoform 3 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30849
R:Bernier, G.; Mathieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
Genomics 38, 19-29, 1996

A:Title: Cloning and characterization of mouse ACF7, a novel member of the dystonin sub
A:Reference number: Z20900; MUID:97124842; PMID:8954775
A:Accession: T30849
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-2033 <BER>
A:Cross-references: UNIPARC:UPI000016D11A; EMBL:U67205; NID:g1675225; PID:g1675226; PIDN
C:Genetics:
A:Gene: mACF7
A:Map position: 4

Query Match 10.6%; Score 84.5; DB 2; Length 2033;
Best Local Similarity 25.4%; Pred. No. 4.1;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMMLAVLIWTGAENLHVKISCLDWLWVSVIPVAESRNLYIFADELHLGMGCPA--- 59
DB 577 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQFVQCE 621

QY 60 NRIHTYVVEFYLVNRDCGIRTVWSEETLLFOTELYFTPRNIDHDPQEIHLCEST-SRK- 117
DB 622 SDVIMYIQCEGLIRQLQVDQLRDEKYYQLEELAFVRMLQDELVTLRLECTNLYRKG 681

QY 118 ---SVMLTPVS--TENEIKLDP 134
DB 682 HFSSLELVPPSTLTTHLKAEP 703

RESULT 5
G70519
hypothetical protein Rv1914c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70519
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70519
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-135 <CON>
A:Cross-references: UNIPROT:O07719; UNIPARC:UPI00000D1020; GB:Z97193; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1914c
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1914c

Query Match 9.8%; Score 78.5; DB 2; Length 135;
Best Local Similarity 34.7%; Pred. No. 0.66;
Matches 17; Conservative 9; Mismatches 14; Indels 9; Gaps 2;

QY 13 IWTGAENLHVKISCLDWLWVSVIPV-----AESRNLYIFADELHLGMG 56
DB 40 LWMGAGSLHVK-----MGWAFADIPLTITKAEATNARVYAGVHFGG 84

RESULT 6
A24699
stearyl-CoA 9-desaturase (EC 1.14.19.1) [similarity] - rat
N:Alternate names: acyl-CoA desaturase; Delta9 fatty acid desaturase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: A24699; JX0150; I58133
R:Thiede, M.A.; Ozols, J.; Strittmatter, P.
J. Biol. Chem. 261, 13230-13235, 1986
A:Title: Construction and sequence of cDNA for rat liver stearyl coenzyme a desaturase.
A:Reference number: A24699; MUID:87008535; PMID:2428815
A:Accession: A24699
A:Molecule type: mRNA

A;Residues: 1-358 <THI>
A;Cross-references: UNIPROT:P07308; UNIPARC:UPI000004210A; GB:J02585; NID:g206859; PIDN:
J;Mihata, K.
J. Biochem 108, 1022-1029, 1990
A;Title: Structure and regulation of rat liver microsomal stearyl-CoA desaturase gene.
A;Reference number: JX0150; MUID:91210202; PMID:1982442
A;Accession: JX0150
A;Molecule type: DNA
A;Residues: 1-290, 'A', 292-358 <MIH>
A;Cross-references: UNIPARC:UPI0000069BF
R;Baba, H.; Fuss, B.; Watson, J.B.; Zane, L.T.; Macklin, W.B.
Neurochem. Res. 19, 1091-1099, 1994
A;Title: Identification of novel mRNAs expressed in oligodendrocytes.
A;Reference number: I58133; MUID:95098214; PMID:7800118
A;Accession: I58133
A;Status: preliminary; translated from GB/EWEL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-4, 'I', 6-10, 'G', 12-13, 'SA', 16-20, 'A', 22-25, 'GQ', 28-30, 'G', 32-33, 'FE', 36, 'N'
A;Cross-references: UNIPARC:UPI000008198; GB:S75730; NID:g861452; PIDN:AAB32826.1; PID:
C;Superfamily: acyl-CoA desaturase; stearyl-CoA desaturase homology
C;Keywords: oxidoreductase; unsaturated fatty acid biosynthesis
F;113-307/Domain: stearyl-CoA desaturase homology <SDH>

Query Match 9.6%; Score 77; DB 1; Length 358;
Best Local Similarity 25.0%; Pred. No. 3.1;
Matches 38; Conservative 22; Mismatches 50; Indels 42; Gaps 8;

QY 42 RNLYIFADELHLG-----MCCPANRIHTYVE-FIYLVRDCGI----- 78
DB 73 RNILMA-LHVGALYGITLIPSSKVYTLGWIFYLLISALGITAGAHRLWSHRTYKARL 131

QY 79 --RTRVSEETLAFQELYTPRN--IDHPQEIHLCESTRKSV-----WLT-----PV 124
DB 132 PLRIFLIANTAFQNDVYEWADHRAHKKFSETHADPHNSRRGFFFSHVGLLVKHPA 191

QY 125 STENEIKLDPSPIAD----FQTTABELGLLI 152
DB 192 VKEKGKLDNSDLKAEKLVMFQRYYPGLLL 223

RESULT 7
A82094
thiamin biosynthesis lipoprotein ApbE VC2289 [imported] - Vibrio cholerae (strain N16961)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82094
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82094
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-367 <HEI>
A;Cross-references: UNIPROT:Q9KPS3; UNIPARC:UPI000000C3262; GB:AE004300; GB:AE003852; NID:
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2289
A;Map position: 1
C;Superfamily: hypothetical protein HI0172

Query Match 9.6%; Score 76.5; DB 2; Length 367;
Best Local Similarity 26.9%; Pred. No. 3.6;
Matches 39; Conservative 21; Mismatches 42; Indels 43; Gaps 9;

QY 19 NLHVKISCLDMLWMSVIPVA-----ESRNLYIFADELHGMGC--PANRIH----- 63
DB 5 NWILKYSMWI-WFRVSVFFPISFIYERSKQVRNWLVALASLLILAGCEKPAEQVHLSGPT 63

QY 64 ---TYVVEFYLVDCGIRTRVSEETLAFQELYTPRNIDHPQEIHLCESTRKSVW 120
DB 111 VYVTSIDYLLIILFLAQLSNQKQWHIYAGQ-YLGTGLLVGASLVAAVYVNVFPEWM 69

DB 64 MGTYYNIKIYQ-----QPGIADSKTL--QTE-----IDRLLEEVDQDMSTYRKDSE 107
QY 121 LTPVSV--TENEIKLDPSPIADFQT 143
DB 108 LSRFNQHTSSE-----PFAVSTQT 126

RESULT 8
AE0152
probable bacteriophage protein YP01246 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE0152
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0152
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-468 <KUR>
A;Cross-references: UNIPROT:Q8ZGPI; UNIPARC:UPI000000DCD86; GB:AL590842; PIDN:CAC90080.1
C;Genetics:
A;Gene: YP01246

Query Match 9.6%; Score 76.5; DB 2; Length 468;
Best Local Similarity 20.4%; Pred. No. 4.8;
Matches 31; Conservative 20; Mismatches 70; Indels 31; Gaps 3;

QY 14 WTCAENLHVKISCLDMLWMSVIPVAESRNLYIFADELHGMGCNPNR--IHTYVVEFY 71
DB 19 WOMSEHLHRASPRGVPFVVVS-----GQTFGRQVTHSYVPYRDT 59

QY 72 LVYDCGIRTRVSEETLAFQELYTPRNIDHPQEIHLCESTRKSVWLTPTVSTENEIK 131
DB 60 YIEDLGNTRKIVLKGILQNSQIYAPDVMQTRQSLAACEMSGPGTLVHPTLGEMTVS 119

QY 132 LDPSPIAD-----FQTTABELGLLI 153
DB 120 ISEAGLLIDDSFSSRVSFSLTAISGLRAF 151

RESULT 9
H95188
cadmium resistance transporter, probable SP1625 [imported] - Streptococcus pneumoniae
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: H95188
R;Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei-
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrisor
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: H95188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-204 <KUR>
A;Cross-references: UNIPROT:Q97PJ0; UNIPROT:Q8DNYO; UNIPARC:UPI0000051908; GB:AE005672;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1625

Query Match 9.5%; Score 76; DB 2; Length 204;
Best Local Similarity 24.7%; Pred. No. 2;
Matches 39; Conservative 27; Mismatches 46; Indels 46; Gaps 9;

QY 22 VKISCLDMLWMSVIPVAE---SRNLYIFADELHGMGC--CPANRIHTYVVEFY----- 69
DB 11 VYVTSIDYLLIILFLAQLSNQKQWHIYAGQ-YLGTGLLVGASLVAAVYVNVFPEWM 69

```

Qy 70 -----IYLVRDCGIRTVSBEILLFQTELYFTPRNIDHDPQEI--HLECSTSRKSV 119
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 70 VGLIAGLPIYL---GIRFAIVGEDA-----EEEEEEIIEEQSKANQLF 111
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy 120 W-----LFPVSTENIEIKLDPSPFIADFOQTABELGLLIF 153
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 112 WTVTLTTIATSGDNLGI-YIPYPASLDWSQTLVALLVF 148
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 10
A98055
hypothetical protein cadb [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A98055
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A98055
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <KUR>
A:Cross-references: UNIPROT:Q97PJ0; UNIPROT:Q8DNY0; UNIPARC:UPI0000051908; GB:AE007317;
C:Genetics.
A:Gene: cadb

```

	Query Match	9.5%	Score 76;	DB 2;	Length 204;
	Best Local Similarity	24.7%	Pred. No. 2;		
	Matches 39;	Conservative 27;	Mismatches 46;	Indels 46;	Gaps 9
Qy	22 VKISCSLDMLVSVIPVAE-----SENLYIFADELHLGMG--CPANRIHTYVYEF-----	69			
Db	11 VYIISTIDYLLIILFPAQLSONKQKWHIYAGQ-YLGTGLLVGASLVAAAYVNVFPEEMM	69			
Qy	70 -----IYLVRCGIRTVSVRETLTFQTELYFTPRNIDHDPOEI--HLECSSTRSRKV	119			
Db	70 VGLLGLPIYL---GIRFAIVGEDA-----EEEEIEITLEREQSKANQLF	111			
Qy	120 W-----LTFVSTENEIKLDPSPFIADFQTTABELGLLIIF	153			
Db	112 WTVTLLITIASGDNGLI-YIPVFASLDMSOTLLVALLVF	148			

RESULT 11

T45895

hypothetical protein F4P12.250 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C:Accession: T45895

R:Blöcker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quétier, F.; Salanoubat, M.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223016

A:Accession: T45895

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-256 <BLO>

A:Cross-references: UNIPROT:Q9LFG4; UNIPARC:UPI00000A485C; EMBL:AL132966

A:Experimental source: cultivar Columbia; EAC clone F4P12

C:Genetics:

A:Map position: 3

A:Introns: 2/3; 53/1; 122/3; 152/3

A:Note: F4P12.250

```

Query Match      9.5%; Score 76; DB 2; Length 256;
Best Local Similarity 20.8%; Pred. No. 2.6;
Matches 36; Conservative 18; Mismatches 51; Indels 68; Gaps 5;

Ov 1 MALEVLMLLAVLIIVTGAENLHVKTSC-----SIDWL-----WVSVI 36

```

1 MALEVLMLLAVLIWTGAENLHVKISC-----SLDWL-----MVSVI 36

[illegible]

RESULT 12

R33001 12
 T33152
 hypothetical protein T04D1.4 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 28-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T33152
 R;Davidson, S.; Wohldmann, P.
 submitted to the EMBL Data Library, May 1998
 A;Description: the sequence of C. elegans cosmid T04D1.
 A;Reference number: Z21292
 A;Accession: T33152
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2957 <DAV>
 A;Cross-references: UNIPROT:O61845; UNIPARC:UPI000007A573; EMBL:AF067617; P
 A;Experimental source: strain Bristol N2; Clone T04D1
 C;Genetics:
 A;Gene: CESP:T04D1.4
 A;Map position: 1
 A;Introns: 122/3; 293/3; 515/3; 1205/2; 1577/1; 2221/3; 2776/1; 2864/3

```
Query Match      9.5%; Score 76; DB 2; Length 2957;
Best Local Similarity 24.%; Pred. No. 50;
Matches 23; Conservative 17; Mismatches 35; Indels 18; Gaps 3;

Qy 62 IHTVY-----EFTLVRCGIRTV---VSEETLLFQTLYLTFPRNIDHPQE 107
    |||||:::||||:::||||:::||||:::|
Db 2056 VHDPVYNKLNLTRESEQEFTVRVSYGVTNGEGNEDTIFKTLSPILEKTDEDCE 2115
    ::::|:::|:::|:::|:::|:::|

Qy 108 ----IHLECSRKSVMLTPVSTENEIKLDPS 136
    ::::|:::|:::|:::|:::|
Db 2116 OLMCVLAMCORAOGSNDLKPIDUKRAMSIDPMP 2148
    ::::|:::|:::|:::|:::|
```

RESULT 13

accession 13
 zeta pellucida glycoprotein C - cat
 S70399
 C:Species: Felis silvestris catus (domestic cat)
 C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
 C:Accession: S70399
 R:Harris, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, A.G.
 DNA Seq. 4, 361-393, 1994
 A:Title: Cloning and characterization of zeta pellucida genes and cDNAs from a varie
 A:Reference number: S70396; MUID:95143578; PMID:7841460
 A:Accession: S70399
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-424 <HAR>
 A:Cross-references: UNIPROT:P48832; UNIPARC:UPI000013C463; EMBL:U05778; NID:g458272;
 C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology
 F:43-299/Domain: ZP domain homology <ZPH>

Query Match	9.4%;	Score 75.5;	DB 2;	Length 424;
Best Local Similarity	22.6%;	Pred. No. 5.4;		
Matches 42;	Conservative	27;	Mismatches 50;	Indels 67;
				Gaps 11;

Qy 1 MALEVLMLLAVLIWCA-----ENLH-----VKISCSDLWLVSVVTPVAESR 422

Db 1 MGLSYGLTICFLWAGTGLCYPTTTDKTHPSLPSSPVVVECRHAWLVNV-----SK 55

Ov 43 NLY-----TFADBLHGMCPAN-----RIHTVVSFYILVROCGIIRTVS 84

Db 56 NLFOTGRLVRADLTG---PENCEPLISGSDDTVRFEVELHK-----CGNSVQ-VT 104
QY 85 EETLLFOTELYFTRNIDH-----DQEIHLCESTRKSVWLTPTVSTENEIKLDPSPFI 138
Db 105 EDALVYSTFLHNPMPGNLSILTRNRAVPFIECRYPRHS-----NVSSEAIL-----PTW 155
QY 139 ADFOTT 144
Db 156 VPRFTT 161

RESULT 14
I58353
LIM protein kinase (EC 2.7.1.-) 1 [similarity] - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 09-Jul-2004
C:Accession: I58353
R:Nunoue, K.; Ohashi, K.; Okano, I.; Mizuno, K.
Oncogene 11, 701-710, 1995
A:Title: LIMK-1 and LIMK-2, two members of a LIM motif-containing protein kinase family.
A:Reference number: I58353; MUID:95380177; PMID:7651734
A:Accession: I58353
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-647 <RES>
A:CROSS-references: UNIPROT:P53669; UNIPARC:UPI000012E690; GB:D31873; NID:g1684611; PTDN
A:Experimental source: strain Wistar; tissue brain
C:Genetics:
A:Gene: limk-1
C:Superfamily: LIM protein kinase; LIM metal-binding repeat homology; protein kinase hom
C:Keywords: ATP; duplication; phosphotransferase; serine/threonine/tyrosine-specific pro
F:25-75/Domain: LIM metal-binding repeat homology <LIM1>
F:84-137/Domain: LIM metal-binding repeat homology <LIM2>
F:337-611/Domain: protein kinase homology <KIN>
F:345-353/Region: protein kinase ATP-binding motif

Query Match 9.4%; Score 75; DB 1; Length 647;
Best Local Similarity 20.1%; Pred. No. 10;
Matches 35; Conservative 15; Mismatches 62; Indels 62; Gaps 4;

QY 21 HVKISCSLWLMVSVIPVAE-----SRNLIFADLHLG 54
Db 129 HSKLYGCQCYQVTVTFVIEQLIPDSGSHLPHVTVLVIPASAHKRGSLSVIDPPHGP 188
QY 55 MGCANRIHT-----YVVEFIYLRDCGIRTVWSEETLLFQTELY 95
Db 189 PGCGTEHSHTVRVQVDPGCMSPDVKNISHIGDRILEINGTPIRNVPDLDEIDLLIQETSR 248
QY 96 FTRPRNDHDPQEIHLCESTRKSVWLTPTVSTENEIKLDPSPFIADFTTABELG 149
Db 249 LLQLTLEHDPHD-----SLHGPPVS-----DPSPLASPVHTPFGQAG 285

RESULT 15
H84028
DNA repair protein BH3032 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: H84028
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: H84028
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <STO>
A:CROSS-references: UNIPROT:Q9K8H4; UNIPARC:UPI00001330B1; GB:AP001517; GB:BA0000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3032
C:Superfamily: DNA repair protein radc

Query Match 9.3%; Score 74.5; DB 2; Length 232;
Best Local Similarity 23.9%; Pred. No. 3.3;
Matches 45; Conservative 27; Mismatches 69; Indels 47; Gaps 10;

QY 2 ALEVLMLAVLIWTGAEN-----LHVKISCSLDWL-----MVSVIPVAESRNLYI 46
Db 28 ALSNQEIVAIMLRTGTGNESVLQAOHVLCFGLRLREATVYBELTSIHGIGEAKE- 86
QY 47 FADLHLGMCPCANRIHTYVVEFIYLRDCGIRTVWSEETLLFQTE-----LYFTPRN-- 100
Db 87 FCAAIELG-----RRITHMQEMDRYVIRTPEDEVSRFVMEEMRFLSQENFVCLYLNTKNQV 141
QY 101 -----IDHDPQEIHLCESTRKSVWLTPTVSTENEIKLDPSPFIADFTT-- 144
Db 142 LHKQTVFISGLNASIVH--PREVPKE--ALRRSA-ASLILHNPSPGDPTSPREDIEVTHR 197
QY 145 ABELGLLI 152
Db 198 LAQVGKLI 205

Search completed: May 9, 2006, 10:15:40
Job time : 18.1 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2006, 10:11:54 ; Search time 264.6 Seconds
(without alignments)
407.959 Million cell updates/sec

Title: US-10-664-025-3903

Perfect score: 799

Sequence: 1 MALEVLMLLAVLIWTGAENL.....PSPFIADFOTTAELGLLIF 153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	789	98.7	158	Q86WS3 HUMAN	Q86ws3 homo sapien
2	785	98.2	158	Q8N9U6 HUMAN	Q8n9u6 homo sapien
3	431.5	54.0	164	Q4FZG8 MOUSE	Q4fzg8 mus musculus
4	234	29.3	173	Q9J183 MOUSE	Q9j183 mus musculus
5	230	28.8	173	Q80WV3 MOUSE	Q80wv3 mus musculus
6	223.5	28.0	212	Q9HBJ0 HUMAN	Q9hbjo homo sapien
7	219.5	27.5	173	Q4V7E2 RAT	Q4v7e2 rattus norv
8	142	17.8	202	Q925U0 MOUSE	Q925u0 mus musculus
9	103	12.9	92	Q8K4N9 MOUSE	Q8k4n9 mus musculus
10	92.5	11.6	422	ZP3 MEGAU	P23491 mesocricetu
11	86.5	10.8	408	Q4T0A6 TETNG	Q4t0a6 tetraodon n
12	85.5	10.7	1062	Q6ZSD7 HUMAN	Q6zsd7 homo sapien
13	85.5	10.7	5430	MACP1_HUMAN	Q9upn3 homo sapien
14	85.5	10.7	5430	Q5VW20 HUMAN	Q5vw20 homo sapien
15	85	10.6	1003	Q95YL0 CIOSEA	Q95yl0 ciona savig
16	85	10.6	1241	Q81GP9 BACOR	Q81gp9 bacillus ce
17	85	10.6	1241	Q6HM43 BACHK	Q6hm43 bacillus th
18	85	10.6	1241	Q63EM2 BACCZ	Q63em2 bacillus ce
19	84.5	10.6	1900	Q5DTZ2 MOUSE	Q5dtz2 mus musculus
20	84.5	10.6	5327	MACP1_MOUSE	Q9gxz0 mus musculus
21	84.5	10.6	7354	Q4PLL5 MOUSE	Q4pll5 mus musculus
22	84	10.5	526	Q6FLR6 CANGA	Q6flr6 candida gla
23	84	10.5	1241	Q73C23 BACC1	Q73c23 bacillus ce
24	84	10.5	1241	Q81TW1 BACAN	Q81tw1 bacillus an
25	82.5	10.3	272	Q8QLN1 NPVMC	Q8qln1 mamestra co
26	82.5	10.3	272	Q71AK8 9NUCL	Q71ak8 mamestra co
27	82.5	10.3	486	Q804D8 BUFAR	Q804d8 bufo arenar
28	82	10.3	442	Q72B14 DESVH	Q72b14 desulfovibr
29	80.5	10.1	2502	Q4RQ55 TETNG	Q4rq55 tetraodon n
30	79.5	9.9	418	Q8K405 9RODE	Q8k405 lagurus leg
31	79.5	9.9	433	Q91191_WSSV	Q91191 white spot

32	79	9.9	1241	2	Q4MPH4 BACCE	Q4mph4 bacillus ce
33	79	9.9	2223	2	Q5CVH3_CRYPV	Q5cvh3 cryptospori
34	78.5	9.8	135	2	Q71ZA9_MYCBO	Q71za9 mycobacteri
35	78.5	9.8	135	2	Q07719_MYCTU	Q07719 mycobacteri
36	78.5	9.8	361	2	Q4T2W2_TETNG	Q4t2w2 tetraodon n
37	78	9.8	501	2	Q9B1D0_ANOGA	Q9b1d0 anopheles g
38	78	9.8	501	2	Q7PEP3 ANOPEL	Q7pep3 anopheles g
39	77.5	9.7	272	2	Q8JME9_9NUCL	Q8jme9 mamestra co
40	77	9.6	358	1	ACOD1_RAT	P07308 rattus norv
41	77	9.6	3032	2	Q60XP7 CAEBR	Q60xp7 caenorhabdi
42	76.5	9.6	253	2	Q6NRU2 XENLA	Q6nru2 xenopus lae
43	76.5	9.6	367	2	Q9KPS3 VIRCH	Q9kps3 vibrio chol
44	76.5	9.6	390	2	Q6C4M0_YARLI	Q6c4m0 yarrowia li
45	76.5	9.6	468	2	Q8ZGPI_YERPE	Q8zgp1 yersinia pe

ALIGNMENTS

RESULT 1
Q86WS3_HUMAN PRELIMINARY; PRT; 158 AA.
AC Q86WS3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein FLJ36198.
GN Name=FLJ36198;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez Y., Bouffard G.G., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A., et al.
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048121; AAH48121.1; -; mRNA.
DR EMBL; BC036256; AAH36256.1; -; mRNA.
DR Ensembl; ENSG00000149507; Homo sapiens.
KW Hypothetical protein.
SQ SEQUENCE 158 AA; 17971 MW; EE843D6FB8AB81FF CRC64;

```
Query Match      98.7%; Score 789; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.1e-73;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEVLMLLAVLWITGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHLMGCPAN 60
   |||||
Db 1 MALEVLMLLAVLWITGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHLMGCPAN 60
   |||||

QY 61 RIHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTSRKSVW 120
   |||||
Db 61 RIHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTSRKSVW 120
   |||||

QY 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151
   |||||
Db 121 LTPVSTENEIKLDPSPFIADFTTAELGLL 151
   |||||

RESULT 3
Q4FZG8_MOUSE PRELIMINARY; PRT; 164 AA.
AC Q4FZG8;
DT 13-SEP-2005 (TremBLrel. 31, Created)
DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TremBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whaley M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC099498; AAH99498.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 18849 MW; 49ADE19216BB8606 CRC64;

Query Match      54.0%; Score 431.5; DB 2; Length 164;
Best Local Similarity 55.9%; Pred. No. 2.2e-36;
Matches 85; Conservative 25; Mismatches 41; Indels 1; Gaps 1;

QY 1 MALEVLMLLAVLWITGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHLMGCPA 59
   |||||
Db 1 MALEVLMLLAVLWITGAENLHVKISCSLDWLMVSVIPVAESRNLYIFADELHLMGCPA 60
   |||||

QY 60 NRHTVYVEFIYLRDCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTSRKSV 119
   :|||
Db 61 TQIHAHQYDFIYPSHCGIRTVSVSEETLLFQTELYFTPRNIDHDQEIHLCESTSRKSV 120
   |||||

QY 120 WLTPVSTENEIKLDPSPFIADFTTAELGLL 151
   |||||
```


QY	7	MLLAVLWITGAENLHVKISCSLDWLMVSVIPVAESRNLVIFADELHGLMGCPANRIHTYV	66
Db	11	VLVAFWFLPSYSEQNVNLCSTDWFMVTHPFLNNDVVHFVHVLGLGCPNHHPRF	70
QY	67	YEFIVLRDCGIRTRVSVSETLLFQTELYFTPRNDHDPQEIHLCEGTSRKSVMWLT-PVS	125
Db	71	YQFHYRVTECGIRIKAVSPDVVISSEIHYASKG-SSTKYVIPVSCAAPRRSPWLTTPKYS	129
QY	126	TE 127	
Db	130	AK 131	
RESULT 5			
Q80W3 MOUSE			
ID	Q80W3_MOUSE	PRELIMINARY;	PRT; 173 AA.
AC	Q80W3		
DT	01-JUN-2003	(Tremblrel. 24, Created)	
DT	01-JUN-2003	(Tremblrel. 24, Last sequence update)	
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)	
DE	Placental specific protein 1.		
GN	Name=Plac1;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C57BL/6J; TISSUE=Embryo;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,		
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C57BL/6J; TISSUE=Embryo;		
RL	Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC051666; AAHS1666.1; -; mRNA.		
DR	MGI; MGI:1926287; Plac1.		
DR	GO; GO:0005615; C:extracellular space; TAS.		
SQ	SEQUENCE 173 AA; 19573 MW; 79DDC0179BD730B0 CRC64;		
Query Match			
Best Local Similarity 28.8%; Score 230; DB 2; Length 173;			
Matches 42; Conservative 32; Mismatches 46; Indels 2; Gaps 2;			
QY	7	MLLAVLWITGAENLHVKISCSLDWLMVSVIPVAESRNLVIFADELHGLMGCPANRIHTYV	66
Db	11	VLVAFWFLPSYSEQNVNLCSTDWFMVTHPFLNNDVVHFVHVLGLGCPNHHPRF	70
QY	67	YEFIVLRDCGIRTRVSVSETLLFQTELYFTPRNDHDPQEIHLCEGTSRKSVMWLT-PVS	125
Db	71	YQFHYRVTECGIRIKAVSPDVVISSEIHYASKG-SSTKYVIPVSCAAPRRSPWLTTPKYS	129

QY	126	TE 127	
Db	130	AK 131	
RESULT 6			
Q9HBJ0 HUMAN			
ID	Q9HBJ0_HUMAN	PRELIMINARY;	PRT; 212 AA.
AC	Q9HBJ0		
DT	01-MAR-2001	(Tremblrel. 16, Created)	
DT	01-MAR-2001	(Tremblrel. 16, Last sequence update)	
DT	13-SEP-2005	(Tremblrel. 31, Last annotation update)	
DE	PLAC1 (Hypothetical protein FLJ90605) (Placenta-specific 1).		
GN	Name=PLAC1;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=20453190; PubMed=10995572; DOI=10.1006/geno.2000.6302;		
RA	Cocchia M., Huber R., Pantano S., Chen E.Y., Ma P., Forabosco A.,		
RA	Ko M.S., Schlessinger D.;		
RT	"PLAC1, an Xq26 gene with placenta-specific expression.";		
RL	Genomics 68:305-312(2000).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Placenta;		
RA	Otsuki T., Ota T., Nishikawa T., Hayashi K., Suzuki Y., Yamamoto J.,		
RA	Wakamatsu A., Kimura K., Sakamoto K., Hatanoto N., Kawai Y., Ishii S.,		
RA	Saito K., Kojima S., Sugiyama T., Ono T., Okano K., Yoshikawa Y.,		
RA	Aotaka S., Sasaki N., Hattori A., Okumura K., Nagai K., Sugano S.,		
RA	Isogai T.;		
RT	"Signal Sequence and Keyword Trap in silico for Selection of Full-		
RT	Length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-		
RT	Capped cDNA Libraries."		
RL	DNA Res. 12:117-126(2005).		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Placenta;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,		
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[4]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Placenta;		
RA	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Placenta;		
RG	NIH MGC Project;		
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF234654; AAG22596.1; -; mRNA.		

QY 71 YLVDCGCGIRTRVWSEETLLFOTELYTFPRNIDHPDQBHLECSRKSVMLT-PVSTE 127
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 75 YRVTEGCGIRAKVPDWVIYSIIHVASKG-SSARYIVIPVSCAPRSPWLTKPSAK 131
| | | | | : | | | | : | | | | : | | | | : | | | | :

RESULT 8

Q925U0 MOUSE
ID Q925U0 MOUSE PRELIMINARY; PRT; 202 AA.
AC Q925U0,
AD
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBurel. 19, Last sequence update)
DT 10-MAY-2005 (TrEMBurel. 30, Last annotation update)
DE Initiate factor 3 (Oocyte-secreted protein 1 precursor) (In vitro
DE fertilized eggs cDNA, RIKEN full-length enriched library,
DE clone:7420700J01 product:INITIATE FACTOR 3, full insert
DE sequence).
GN Names:Oospi; Synonyms=if3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MANO H., Nakatani S., Aoyagi R., Iehii R., Iwai Y., Shimoda N.,
RA Jincho Y., Hiura H., Hirose M., Mochizuki C., Yuri M., Im R-H.,
RA Funada-Wada U., Wada M.;
RT "If3, a novel cell-differentiation factor, highly expressed in the
RT murine liver and ovary.";
RL Biochem. Biophys. Res. Commun. 297:323-328(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22222142; PubMed=122377121; DOI=10.1016/S0006-291X(02)02194-0;
RA Mano H., Nakatani S., Aoyagi R., Iehii R., Iwai Y., Shimoda N.,
RA Jincho Y., Hiura H., Hirose M., Mochizuki C., Yuri M., Im R-H.,
RA Funada-Wada U., Wada M.;
RT "If3, a novel cell-differentiation factor, highly expressed in the
RT murine liver and ovary.";
RL Biochem. Biophys. Res. Commun. 297:323-328(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21610718; PubMed=11747200; DOI=10.1002/gene.10010;
RA Yan C., Pendola F.L., Jacob R., Lau A.B., Epig J.J., Matzuk M.M.;
RT "Oosp1 encodes a novel mouse oocyte-secreted protein.";
RA Genesis 31:105-110(2001).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RA Meth. Enzymol. 303:19-44(1999).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyoisawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Hoffell D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=In vitro fertilized eggs;
MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266.

OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Cricetidae; Cricetinae; Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC MEDLINE=91078540; PubMed=2257975;
RX Kinloch R.A., Ruiz-Seller B., Wasserman P.M.;
RA "Genomic organization and polypeptide primary structure of zona
RT pellucida glycoprotein h2p3, the hamster sperm receptor.";
RL Dev. Biol. 142:414-421(1990).
CC -!- FUNCTION: The mammalian zona pellucida, which mediates species-
CC specific sperm binding, induction of the acrosome reaction and
CC prevents post-fertilization polyspermy, is composed of three to
CC four glycoproteins, ZP1, ZP2, ZP3, and ZP4. ZP3 is essential for
CC sperm binding and zona matrix formation.
CC -!- SUBUNIT: Polymers of ZP2 and ZP3 organized into long filaments
CC cross-linked by ZP1 homodimers (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
CC matrix upon proteolytic cleavage.
CC -!- TISSUE SPECIFICITY: Oocytes.
CC -!- DEVELOPMENTAL STAGE: Growing oocytes.
CC -!- DOMAIN: The ZP domain is involved in the polymerization of the ZP
CC proteins to form the zona pellucida.
CC -!- PTM: Proteolytically cleaved before the transmembrane segment to
CC yield the secreted ectodomain incorporated in the zona pellucida.
CC -!- PTM: N-glycosylated (By similarity).
CC -!- PTM: O-glycosylated; removal of O-linked glycans may play an
CC important role in the postfertilization block to polyspermy (By
CC similarity).
CC -!- SIMILARITY: Belongs to the ZP domain family. ZPC subfamily.
CC -!- SIMILARITY: Contains 1 ZP domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M63629; AAA37079.1; -; mRNA.
CC InterPro; IPR001507; Endoglin/CD105.
CC Pfam; PF00100; Zona pellucida; 1.
CC PRINTS; PR00023; ZP_ELUUCIDA.
CC SMART; SM00241; ZP; 1.
CC PROSITE; PS00682; ZP_1; 1.
CC PROSITE; PS51034; ZP_2; 1.
CC Cleavage on pair of basic residues; Extracellular matrix; Receptor;
KW Fertilization; Glycoprotein; Pyrrolidone carboxylic acid; Receptor;
KW Signal; Transmembrane.
FT SIGNAL 1 22
FT CHAIN 23 349
FT PROPEP 350 422
FT TOPO_DOM 23 386
FT TRANSMEM 387 407
FT TOPO_DOM 408 422
FT DOMAIN 45 306
FT COMPBIAS 119 158
FT COMPBIAS 208 257
FT MOD_RES 23 23
FT
FT CARBOHYD 32 32
FT CARBOHYD 34 34
FT CARBOHYD 146 146
FT CARBOHYD 155 155
FT CARBOHYD 161 161
FT CARBOHYD 162 162
FT CARBOHYD 271 271
FT CARBOHYD 302 302
FT CARBOHYD 78 98
FT DISULFID 216 281
FT DISULFID 238 299

SQ SEQUENCE 422 AA; 45827 MW; D0F95BE7FF8E7E01 CRC64;
Query Match 11.6%; Score 92.5; DB 1; Length 422;
Best Local Similarity 23.4%; Pred. No. 0.82;
Matches 51; Conservative 30; Mismatches 66; Indels 71; Gaps 12;

QY 1 MALEVLMLAVLIWTGAENLHVKISCSLD-WLMVSVIP-----VAE 40
Db 1 MGLSVQLLCLLLCGA-----KQCCSQPLMLLPGGTPTPGKLTSSVEVECLAEVLVTV 55
QY 41 SRNLY-----IFADELHLCM-OC-PANRHTVYVEFYLVRCGIRTVVSEETLLFQTE 93
Db 56 SRDLFTGKLIQEDLTGSENCRLPVSVDVRFKAQLHECSNRVQ-VTEALVYSTV 114
QY 94 LVFTPRNID-----HDPQEIHLCESTR-----KSVML---TPVSTENEI----- 130
Db 115 LLHQPRPVPGLSILRTNRADVPICRYPQGNVSSHAIPTWVPFTTSSSEKLVFSLR 174
QY 131 -----KLDPSPIADFQ-----TTAEELGLLIF 153
Db 175 LMENWNTEKLSFTSHLGEVAYLOAEVQTGSHLPLLLF 212

RESULT 11
Q4TOA6_TETNG
ID Q4TOA6_TETNG PRELIMINARY; PRT; 408 AA.
AC Q4TOA6;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Chromosome 3 SCAF11271, whole genome shotgun sequence.
GN ORFNames=GSTENG0009437001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC EMBL; CAEE01011271; CAF93676.1; -; Genomic DNA.
SQ SEQUENCE 408 AA; 44840 MW; CABA9142A7302BD2 CRC64;

Query Match 10.8%; Score 86.5; DB 2; Length 408;
Best Local Similarity 30.5%; Pred. No. 3.3;
Matches 39; Conservative 16; Mismatches 38; Indels 35; Gaps 8;

QY 46 IFADELHLCMGCPANRIHTYV-----EPIYLV--RDCGIRTVVSEETLLFOTELY 95
Db 8 VYSDELGV---EQRDCHRAHSEGGEEETILVGLADCGTK-HWSEDAIYTNLLI 63

```

QY 96 FTPT-----NIDHPQEIHLCESTRK-----SVWLTPVSTENEIKLDPSPFIAD----FOT 143
Db 64 FSPRTVPTDGLIRMDIAVIEQCYERKYSLSSSLTPTWV-----PFMATQAAVET 114
QY 144 TABELGLL 151
Db 115 LAFDRLV 122

RESULT 12
Q6ZSD7 HUMAN
ID Q6ZSD7 HUMAN PRELIMINARY; PRT: 1062 AA.
AC Q6ZSD7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FL045612.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Thalamus;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musahino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagehara K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK127519; BC87016.1; -; mRNA.
DR HSSP; Q01082; IAA2.
DR SMR; Q6ZSD7; 341-461.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR SMART; SM00150; SPEC; 2.
DR PROSITE; PS00021; CH; 1.
SQ SEQUENCE 1062 AA; 119927 MW; 243B3FFF7E87F30E CRC64;

Query Match 10.7%; Score 85.5; DB 2; Length 1062;
Best Local Similarity 25.4%; Pred. No. 13;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVIMLLAVLIWTGAENLHVKISCLDWLMVSVIPVAESRNLYIFADELHLMGCPA--- 59
Db 586 LELLQIANKIQNGALNCEKLTAKNTLIQ-----ADAHLESGQPVQCE 630
QY 60 NRIHYVVEFYLVKDCGTRTWVSEETLLFOTELVFTPRNIDHPDQEIHLCEST-SRK- 117
Db 631 SDVMIVYQCEGLIQLQVDLIQLDENYQYQLEELAFVRWLQDELVTLRLECTNLYRKG 690
QY 118 ---SVWLTPVS--TENEIKLDP 134
Db 691 HFTSLVLPSPSTLTTHLKAEP 712

RESULT 13
MACF1_HUMAN
ID MACF1_HUMAN STANDARD; PRT: 5430 AA.
AC Q9UPN3; O75053; Q8WXV2; Q9H540; Q9UKP0; Q9ULG9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Microtubule-actin crosslinking factor 1, isoforms 1/2/3/5 (Actin
DE cross-linking family protein 7) (Macrophilin 1) (Trabeculin-alpha) (620
DE kDa actin-binding protein) (ABP620).

```

```

GN Name=MACF1; Synonyms=ABP620, ACF7, KIAA0465, KIAA1251;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
RX MEDLINE=20001959; PubMed=10529403; DOI=10.1006/bbrc.1999.1538;
RA Okuda T., Matsuda S., Nakatsugawa S., Ichigotani Y., Iwanashi N.,
RA Takahashi M., Ishigaki T., Hamaguchi M.;
RT "Molecular cloning of macrophin, a human homologue of Drosophila
RT kakapo with a close structural similarity to plectin and dystrophin.";
RL Biochem. Biophys. Res. Commun. 264:568-574(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=20026884; PubMed=10559237; DOI=10.1074/jbc.274.47.33522;
RA Sun Y., Zhang J., Kraeft S.-K., Auclair D., Chang M.-S., Liu Y.,
RA Sutherland R., Salgia R., Griffin J.D., Ferland L.H., Chen L.B.;
RT "Molecular cloning and characterization of human trabeculin-alpha, a
RT giant protein defining a new family of actin-binding proteins.";
RL J. Biol. Chem. 274:33522-33530(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 182-4812 (ISOFORM 2), AND ALTERNATIVE SPLICING.
RX MEDLINE=21833812; PubMed=11845288; DOI=10.1007/s00335-001-3037-3;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT "MACF1 gene structure: a hybrid of plectin and dystrophin.";
RL Mamm. Genome 12:852-861(2001).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 868-2350 (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hiroseawa M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RL for large proteins in vitro.";
RN DNA Res. 6:337-345(1999).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1544-5057 (ISOFORM 2).
RA Corby N.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 868-5430 (ISOFORM 5).
RC TISSUE=Brain;
RX MEDLINE=98116662; PubMed=9455484;
RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
RA Nakajima D., Nomura N., Ohara O.;
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
RT from human brain.";
RL DNA Res. 4:345-349(1997).
RN [7]
RP SEQUENCE REVISION.
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
RN [8]
RP SEQUENCE REVISION.
RA Ohara O.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: F-actin-binding protein which may play a role in cross-
CC linking actin to other cytoskeletal proteins. Also binds to
CC microtubules (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=2;
CC IsoId=Q9UPN3-2; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UPN3-1; Sequence=VSP_007341, VSP_012090;
CC Name=3;

```

```

CC      IsoId=Q9UPN3-3; Sequence=Not described;
CC      Name=4;
CC      IsoId=Q96PK2-1; Sequence=External;
CC      Name=5;
CC      IsoId=Q9UPN3-4; Sequence=VSP_012090, VSP_012091;
CC      Note=No experimental confirmation available;
CC      -!- TISSUE SPECIFICITY: Ubiquitously expressed;
CC      -!- SIMILARITY: Belongs to the plakins or cytolinker family.
CC      -!- SIMILARITY: Contains 1 actin-binding domain.
CC      -!- SIMILARITY: Contains 2 CH (calponin-homology) domains.
CC      -!- SIMILARITY: Contains 2 EF-hand domains.
CC      -!- SIMILARITY: Contains 1 SH3 domain.
CC      -!- SIMILARITY: Contains 37 spectrin repeats.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      EMBL; AB029290; BAA83821.1; -; mRNA.
CC      EMBL; AF141968; AAF06360.1; -; mRNA.
CC      EMBL; AF325341; AAL39000.1; -; Genomic DNA.
CC      EMBL; AF325330; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325331; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325332; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325333; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325334; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325335; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325336; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325339; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AF325340; AAL39000.1; JOINED; Genomic DNA.
CC      EMBL; AB033077; BAA86565.1; -; mRNA.
CC      EMBL; AL137853; CAC15920.1; -; Genomic DNA.
CC      EMBL; AB007934; BAA32310.3; -; mRNA.
CC      PIR; T00079; T00079.
CC      HSSP; Q15149; 1MB8.
CC      SNR; Q9UPN3; 68-301.
CC      Ensembl; ENSG00000127603; Homo sapiens.
CC      HGNC; HGNC:13664; MACF1.
CC      MIM; 608271; -.
CC      GO; GO:0005737; C:Cytoplasm; IDA.
CC      GO; GO:0005856; C:cytoskeleton; NAS.
CC      GO; GO:0051015; F:actin filament binding; NAS.
CC      GO; GO:0005509; F:calcium ion binding; NAS.
CC      GO; GO:008017; F:microtubule binding; NAS.
CC      InterPro; IPR001589; Actinin actin bd.
CC      InterPro; IPR001715; Calponin act_bd.
CC      InterPro; IPR011992; EF-Hand type.
CC      InterPro; IPR002048; EF-hand_Ca_bd.
CC      InterPro; IPR003108; GAS2.
CC      InterPro; IPR001452; SH3.
CC      Pfam; PF00307; CH; 2.
CC      Pfam; PF00036; ehand; 2.
CC      Pfam; PF02187; GAS2; 1.
CC      Pfam; PF00435; Spectrin; 27.
CC      ProDom; PD000012; EF-hand; 1.
CC      SMART; SM00033; CH; 2.
CC      SMART; SM00054; EFh; 2.
CC      SMART; SM00243; GAS2; 1.
CC      SMART; SM00150; SPEC; 33.
CC      PROSITE; PS00019; ACTININ_1; 1.
CC      PROSITE; PS00020; ACTININ_2; 1.
CC      PROSITE; PS50021; CH; 2.
CC      PROSITE; PS00018; EF_HAND_1; 2.
CC      PROSITE; PS00222; EF_HAND_2; 2.
CC      PROSITE; PS50002; SH3; FALSE NEG.
CC      Actin-binding; Alternative splicing; Calcium; Coiled coil;
CC      Cytoskeleton; Repeat; SH3 domain.
CC      DOMAIN 1 295 Actin-binding.
CC      DOMAIN 78 181 CH 1.
CC      DOMAIN 194 295 CH 2.

```

```

FT REPEAT 314 355 Spectrin 1.
FT REPEAT 591 623 Spectrin 2.
FT REPEAT 680 784 Spectrin 3.
FT REPEAT 786 800 Spectrin 4.
FT DOMAIN 871 923 SH3.
FT REPEAT 1250 1272 Spectrin 5.
FT REPEAT 1287 1342 Spectrin 6.
FT REPEAT 1455 1534 Spectrin 7.
FT REPEAT 1547 1659 Spectrin 8.
FT REPEAT 1815 1891 Spectrin 9.
FT REPEAT 1932 2042 Spectrin 10.
FT REPEAT 2260 2280 Spectrin 11.
FT REPEAT 2372 2395 Spectrin 12.
FT REPEAT 2398 2507 Spectrin 13.
FT REPEAT 2510 2618 Spectrin 14.
FT REPEAT 2621 2728 Spectrin 15.
FT REPEAT 2731 2838 Spectrin 16.
FT REPEAT 2841 2945 Spectrin 17.
FT REPEAT 2987 3024 Spectrin 18.
FT REPEAT 3136 3163 Spectrin 19.
FT REPEAT 3187 3274 Spectrin 20.
FT REPEAT 3277 3383 Spectrin 21.
FT REPEAT 3386 3492 Spectrin 22.
FT REPEAT 3495 3601 Spectrin 23.
FT REPEAT 3604 3673 Spectrin 24.
FT REPEAT 3713 3819 Spectrin 25.
FT REPEAT 3832 3927 Spectrin 26.
FT REPEAT 3982 4043 Spectrin 27.
FT REPEAT 4046 4152 Spectrin 28.
FT REPEAT 4155 4262 Spectrin 29.
FT REPEAT 4265 4371 Spectrin 30.
FT REPEAT 4374 4481 Spectrin 31.
FT REPEAT 4484 4590 Spectrin 32.
FT REPEAT 4593 4700 Spectrin 33.
FT REPEAT 4707 4808 Spectrin 34.
FT REPEAT 4811 4917 Spectrin 35.
FT REPEAT 4920 4985 Spectrin 36.
FT REPEAT 5030 5054 Spectrin 37.
FT DOMAIN 5083 5118 EF-hand 1.
FT DOMAIN 5119 5154 EF-hand 2.
FT CA_BIND 5096 5107 1 (Potential).
FT CA_BIND 5132 5143 2 (Potential).
FT REGION 5355 5370 4 X 4 AA tandem repeats of [GS]-S-R-[AR].
FT COILED 243 265 Potential.
FT COILED 477 529 Potential.
FT COILED 722 751 Potential.

Query Match 10.7%; Score 85.5; DB 1; Length 5430;
Best Local Similarity 25.4%; Pred. No. 89;
Matches 36; Conservative 23; Mismatches 58; Indels 25; Gaps 5;

QY 3 LEVLMALLAVLWITGAENLHVKISCSLDLWMSVIPVAESRNLYIFADELHLGMCFA--- 59
426 LELLQIANKIQNGALNCEKLTAKNTLQ-----ADAAHLESGQPQVCE 470

QY 60 NRIHTVYVEFYLVDRDCGIRTVVSEITLLFOTELYFTPRNIDHDPQEIHLEST-SRK- 117
471 SDVIMYIQCEGLIRQLQVDLQILRDENYQYLEELAFRVMLQDELVTLRLCTNLYRKG 530

QY 118 ---SVWLTPVS--TENEIKLDP 134
531 HFTSLVLPSTLTTHLKAEP 552

RESULT 14
Q5VW20_HUMAN
ID Q5VW20_HUMAN PRELIMINARY; PRT; 5430 AA.
AC Q5VW20;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Microtubule-actin crosslinking factor 1.
GN Name=MACF1; ORFNames=RP11-69E11.1-001;

```


Job time : 266.6 secs

This Page Blank (uspto)